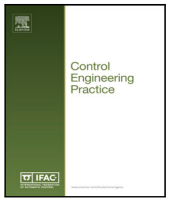
Control Engineering Practice 139 (2023) 105621

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac



Optimal physical human–robot collaborative controller with user-centric tuning

Loris Roveda [a,*], Lorenzo Mantovani [b], Marco Maccarini [a], Francesco Braghin [b], Dario Piga [a]

[a] Istituto Dalle Molle di studi sull'Intelligenza Artificiale (IDSIA), Scuola Universitaria Professionale della Svizzera Italiana (SUPSI), Università della Svizzera Italiana (USI) IDSIA-SUPSI, 6928 Manno, Switzerland

[b] Politecnico di Milano, Department of Mechanical Engineering, Milano, Italy

ARTICLE INFO

Keywords:
Human–robot collaboration
Optimal control
Passive velocity field control
Impedance control
Intelligent control
Preference-based optimization

ABSTRACT

Collaborative robots are increasingly used in different fields of application to physically interact with the human (*e.g.*, manufacturing, rehabilitation, etc.) In order to improve the physical collaboration performance, the interaction between the user and the robot must be safe, smooth, and intuitive. Indeed, this paper proposes a controller which is composed of three nested loops. A Passive Velocity Field Control (PVFC) defines the lowest control layer, ensuring the passivity of the system. The intermediate control layer is defined by a Cartesian impedance controller, managing the interaction between the user and the robot and sending to the PVFC a reference position. The outer layer is defined by a Linear Quadratic Regulator (LQR), detecting the intention of motion of the user and deforming accordingly the Cartesian impedance setpoint to follow such a motion. In addition, to enhance the collaboration performance for each user, a preference-based optimization approach is employed to tune the control parameters, implementing a human-centric tuning procedure. In such a way, the controller is customized for the specific user to establish the proper interaction with the robot. The proposed controller has been evaluated by making use of a Franka EMIKA panda robot as a test platform, comparing the achieved performance with a controller previously developed by some of the authors in a free-motion collaborative task along the z vertical direction. Achieved results show the improved performance obtained by the proposed controller. In addition, an assembly task has also been optimized to show the potential of the proposed control framework in complex and realistic situations.

1. Introduction

Collaborative robotics plays a key role (Makris, 2021) in the emerging Industry 5.0 paradigm (Demir, Döven, & Sezen, 2019), in which the human is put at the center of the production environment, having the robotic systems assisting the operator and executing repetitive and onerous applications (Vicentini et al., 2020). Physical human–robot collaboration (pHRC) is, indeed, one of the most investigated topics (Galin & Meshcheryakov, 2020), making it possible to have the human in a leading position while collaborating with the robot (Roveda, Haghshenas, et al., 2018). Many open issues in the state of the art are still to be overcome, in particular considering safety/stability guarantees in the human–robot interaction while implementing a reactive robot controller, making it possible to detect the intention of motion of the human during the collaboration.

To tackle the above-mentioned issues within the pHRC scenario, this paper proposes a three-layers controller. The inner layer is designed based on the Passive Velocity Field Control (PVFC) (Li & Horowitz, 1999, 2001a, 2001b). The PVFC allows for ensuring the stability of the controlled system, tracking a joint position reference. The intermediate control layer is designed based on the Cartesian impedance controller (Hogan, 1984), providing a controlled compliant behavior to the robot in interaction with the user. The Cartesian impedance control sends the reference to the PVFC (by properly applying the inverse kinematics transformation). The outer layer is designed based on a Linear Quadratic Regulator (LQR) (Bemporad, Morari, Dua, & Pistikopoulos, 2002). The LQR detects the intention of motion of the user, accordingly deforming the Cartesian impedance setpoint to follow such a motion. The LQR is designed to make use of a simplified human–robot interaction dynamics modeling (Artemiadis, Katsiaris, Liarokapis, & Kyriakopoulos, 2010), in which the human's arm is simply modeled as an elastic element (*i.e.*, as a spring). A robust term is included to compute the solution of the LQR in order to consider the variability of the human's arm stiffness (Roveda & Piga, 2020). Furthermore, to properly tune the control gains (*i.e.*, the stiffness and damping of the impedance control, together with the gain of the LQR), a preference-based optimization (PBO) approach has been employed (Bemporad & Piga, 2021). Such a methodology allows to optimize the collaborative

* Corresponding author.
E-mail address: loris.roveda@idsia.ch (L. Roveda).

https://doi.org/10.1016/j.conengprac.2023.105621
Received 14 December 2022; Received in revised form 12 June 2023; Accepted 1 July 2023
Available online 22 July 2023

behavior of the robot based on the specific user's (even implicit) needs. In fact, the user has only to provide qualitative feedback, *i.e.*, giving a preference on the collaborative performance of two sets of control gains. In such a way, no quantitative performance indexes are needed (*i.e.*, no measurements are required), having the optimization able to embed the perception of the user of the collaboration with the robot, thus implementing a human-centric tuning procedure. The proposed approach, indeed, provides a safe and intuitive controller capable to establish an optimized controlled interaction with the human.

In the following Section, the state of the art related to the pHRC control is addressed, to highlight the open issues in the field and the solutions provided by the proposed approach.

1.1. Related work

1.1.1. Physical human–robot interaction

Among other strategies (Magrini & De Luca, 2016; Martinez, Lawson, Durrough, & Goldfarb, 2018), physical human–robot collaboration (pHRC) is commonly enabled by implementing a low-level impedance controller (Hogan, 1984), that provides the robot with a safe and compliant behavior, suitable for interacting with the surrounding environment (including humans Roveda, 2018). The impedance control parameters (*i.e.*, mass/inertia, stiffness, damping, and setpoint) are then tuned/adapted by means of high-level control strategies during the execution of a task (Khan, Herrmann, Al Grafi, Pipe, & Melhuish, 2014), *e.g.*, to achieve human-like adaptability skills (Liang et al., 2014; Yang, Zeng, Fang, He, & Li, 2018), to maximize the human–robot collaboration performance (Kim, Peternel, Lorenzini, Babič, & Ajoudani, 2021), etc.

Machine learning (ML)-based approaches have been investigated to implement such flexible and adaptive controllers. Two types of ML-based solutions are available in the state of the art: model-based ML approaches (Roveda, Maskani, et al., 2020), and model-free ML approaches (Cremer, Das, Wijayasinghe, Popa, & Lewis, 2019). Model-based ML approaches provide powerful algorithms for control tuning purposes that are capable of capturing complex and uncertain interaction dynamics. The main drawback of such strategies consists of the limited variety of task conditions that can be faced by the proposed controllers. In order to be effective, the adopted models should accurately represent the target scenario, losing generalizability. Model-free approaches, on the other hand, allow for achieving acceptable results in a wide set of scenarios by exploiting autonomous tuning through trial and error. However, the tuning procedures are costly (both in terms of the computational resources and the time), requiring a vast amount of trials to achieve the target performance (Roveda, Haghshenas, Caimmi, Pedrocchi, & Molinari Tosatti, 2019). Many efforts are, therefore, put into the development of combined solutions exploiting the advantages of both model-based and model-free ML solutions, such as the ones provided in Dimeas and Aspragathos (2015), Kukker and Sharma (2021), Li, Zhang, Xia, Xie, and Zhu (2018), Medina, Börner, Endo, and Hirche (2019) and Peternel, Tsagarakis, and Ajoudani (2017). However, computational issues are commonly displayed by such approaches, limiting their applicability in real applications. Stability is another issue to be considered in the design of human–robot collaborative controllers (Freeman & Kokotovic, 1996; Grüne & Pannek, 2017; Ma, Kolathaya, Ambrose, Hubicki, & Ames, 2017; Zhang, Li, & Zheng, 2020), which is difficult to be guaranteed for ML-based controllers.

Controllers exploiting the physical interaction dynamics modeling represent another solution for the design and implementation of human–robot collaborative strategies (Hentout, Aouache, Maoudj, & Akli, 2019; Losey, McDonald, Battaglia, & O'Malley, 2018; Rodríguez-Guerra, Sorrosal, Cabanes, & Calleja, 2021). In this case, the human's dynamics becomes critical to be modeled in order to design a robust a stable controller (Mobasser & Hashtrudi-Zaad, 2006; Tee, Burdet, Chew, & Milner, 2004). Indeed, safety is still one of the main issues in the design of such controllers (Zacharaki, Kostavelis, Gasteratos, & Dokas, 2020), especially when high bandwidth is required. In addition, such controllers commonly show a complex design, that is characterized by a high number of control parameters/gains to be tuned (Fu & Zhao, 2020; Gaz, Magrini, & De Luca, 2018; He, Xue, Yu, Li, & Yang, 2020), making it difficult to optimize the robot behavior to maximize the collaborative performance with the human.

A control strategy capable to address the above-mentioned limitations of the state of the art (*i.e.*, reducing computational cost, stability, robustness, and simplified gains tuning) is demanded the implementation of feasible and effective physical human–robot collaborative controllers.

1.1.2. Preference-based optimization

Preference-based optimization (PBO) is a powerful tool to solve black-box problems, in which the modeling of the process is not available/difficult to be accurately described. In addition, such an optimization approach provides an intuitive procedure to deal with processes not allowing quantitative data gathering and/or the definition of performance indexes to be used for optimization purposes. Indeed, researchers have put their attention on optimization methods to minimize black-box functions using preferences, such as employing Particle Swarm Optimization (PSO, Kennedy, 2010; Vaz & Vicente, 2007). Such an optimization drives the evolution of particles only based on the outcome of comparisons between the function values. Preference-based optimization has been also used in combination with Reinforcement Learning (RL) (Lee, Smith, Dragan, & Abbeel, 2021). "Active preference learning" has been also investigated, in which the user iteratively expresses a preference. To this end, Busa-Fekete, Hüllermeier, and Mesaoudi-Paul (2018) presents a review of different active learning algorithms, such as the one developed in Yue, Broder, Kleinberg, and Joachims (2012). An important concept of PBO is represented by RBF functions, that have been studied in Bemporad and Piga (2021) and McDonald, Grantham, Tabor, and Murphy (2007).

The preference-based optimization (PBO) technique has been employed only in a few robotic use cases. User preferences have been combined with demonstrations to reduce the required number of queries (Palan, Landolfi, Shevchuk, & Sadigh, 2019), where a robotic manipulator is trained to reach a goal configuration while avoiding an obstacle. Calibration of model predictive control parameters based on user's preferences is discussed in Zhu, Bemporad, and Piga (2020). In Choi et al. (2020), a Bayesian deep learning method is proposed to optimize the parameters for a navigation task using humans' preference evaluations. In Roveda et al. (2021), PBO has been used to optimize a robotic sealant material deposition task.

No applications of PBO can be found w.r.t. physical human–robot collaboration, which represents one of the most interesting fields of application due to the subjective perception of the human–robot interaction.

1.2. Paper contribution

This paper aims to design a controller to deal with physical human–robot collaboration to face the following state-of-the-art limitations:

- computational effort;
- intrinsic stability of the controller;
- robustness w.r.t. the human–robot coupled modeled dynamics;
- control gains tuning for human assistance optimization.

The proposed controller is composed of three nested loops:

- inner layer: Passive Velocity Field Control (PVFC) loop for joint reference tracking;
- intermediate layer: Cartesian impedance loop for compliance control;
- outer layer: Linear Quadratic Regulator (LQR) for human's intention detection and tracking.

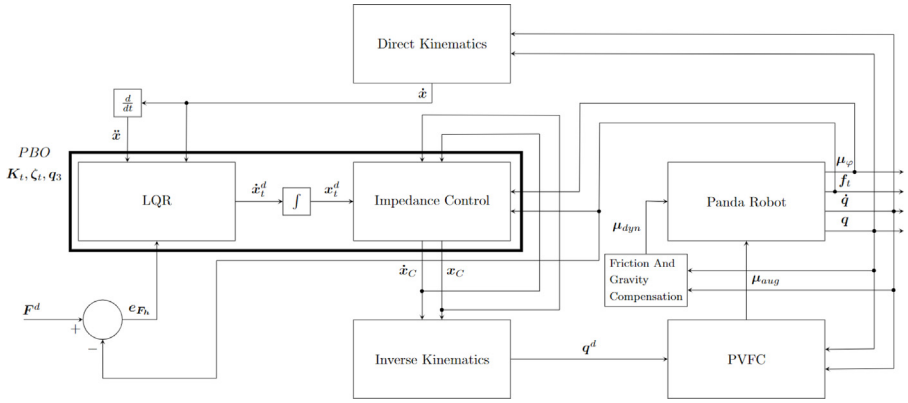


Fig. 1. Control schema highlighting the three control layers: the PVFC, the Cartesian impedance control, and the LQR. The LQR, based on the interaction state, computes the Cartesian impedance setpoint. The Cartesian impedance controller, based on the provided setpoint and on the applied control gains (*i.e.*, mass, damping, and stiffness parameters), computes the reference Cartesian position to be tracked by the low-level PVFC. Such a reference position, taken to the joint space, is indeed used as an input by the PVFC, ensuring the motion tracking for the robot. The PBO for the control gains tuning is also highlighted, optimizing the human's assistance control gains to maximize the collaboration performance.

In such a way, it is possible to:

- implement a real-time controller with a reduced computational cost (*i.e.*, capable to compute the control signals given a high-frequency control loop): employing a simplified human–robot interaction dynamics modeling (Artemiadis et al., 2010), it is possible to analytically derive the LQR gain matrix for its online implementation and computation;
- provide intrinsic stability of the controller at all levels: all the control loops are intrinsically stable (Bemporad et al., 2002; Hogan, 1984; Li & Horowitz, 1999) and the proposed LQR is demonstrated to be stable for the considered human–robot interaction dynamics;
- design a robust controller: the LQR is designed to be robust w.r.t. the human's dynamics modeling (Roveda & Piga, 2020).

In addition, in order to simplify the tuning procedure of the control gains providing assistance to the human during the collaboration with the robot, a human's qualitative feedback-based optimization algorithm is employed. In such a way, the human's (even intrinsic) needs can be catched, maximizing the performance of the collaboration with the robot. To do so, the preference-based optimization (PBO) approach is used (Bemporad & Piga, 2021), implementing a simplified and intuitive human-centric control gains tuning procedure.

Remark 1. It has to be underlined that the PBO is employed only for the tuning of the control gains providing assistance to the human. The other control gains (*e.g.*, the ones related to trajectory tracking) are tuned by means of standard procedures since they do not affect the human–robot collaboration.

The proposed controller has been tested on a Franka EMIKA panda robot for validation, comparing the achieved performance w.r.t. the controller in Roveda, Maskani, et al. (2020) in a free-motion collaborative task along the z vertical direction. 10 volunteers have been involved in the experimental tests, judging the performance of both controllers. The achieved results show the improved performance of the designed controller. In addition, the volunteers evaluated the preference-based tuning procedure, highlighting its usefulness and easiness of use. Finally, an assembly task has also been optimized to show the potential of the proposed control framework in complex and realistic situations.

To summarize, the main contributions of the proposed paper are:

- the design of a stable control framework for human–robot collaboration, being robust w.r.t. the human's arm dynamics. While some works are already providing cascade controllers for interaction tasks (Franceschi, Pedrocchi, & Beschi, 2022; Khoramshahi & Billard, 2020; Mariotti, Magrini, & De Luca, 2019; Tian, Sun, Bajcsy, Tomizuka, & Dragan, 2022), there is still a lack of methodologies addressing together the mentioned issues related to computational effort, intrinsic stability of the controller, and robustness, and this paper is providing a novel solution, combining existing techniques to design a new control framework;
- the application of the PBO for the tuning of the human's assistance gains to maximize the collaboration performance. No contributions in the state of the art can be found where the PBO is employed for the tuning of collaborative robot controllers. Indeed, this contribution is the first exploring the potential of the PBO in this application;
- the extensive testing of the proposed methodologies on a real manipulator to highlight the stability, robustness, and improved performance of the proposed control framework.

1.3. Paper outline

The paper is structured as it follows: Section 2 proposes the complete description of the proposed control schema and its stability analysis, Section 3 describes the preference-based optimization for the tuning of the control gains, Section 4 provides the achieved results, and Section 6 states the conclusions of this work.

2. Controller design

This Section describes in detail the design of the proposed three-layers controller, which is composed of the Passive Velocity Field Control (inner layer), the Cartesian impedance control (intermediate layer), and by the Linear Quadratic Regulator (outer layer). Fig. 1 shows the complete control architecture, highlighting all three control blocks and their connections. In addition, the connection of the control approach with the PBO for human-centric tuning is highlighted. In the following, each of the control layers is described, followed by the stability analysis of the whole control architecture.

2.1. Passive velocity field control

Based on the work in Li and Horowitz (1999, 2001a, 2001b), the PVFC can be designed and employed as the inner loop for the proposed controller (Fig. 1). The main aim of this control loop is to provide passivity guarantees to the controlled system w.r.t. the interaction wrench. To do so, a passive velocity field is designed to control the motion of the manipulator. Taking into consideration the rigid links/joints robot dynamics (Caccavale, Natale, Siciliano, & Villani, 1999) (having that the gravitational term $\mathbf{g}(\mathbf{q})$ and the joint friction effects $\mathbf{f}(\mathbf{q},\dot{\mathbf{q}})$ are already compensated in the robot control law by the control action $\boldsymbol{\mu}_{dyn}$ as in Roveda, Bussolan, Braghin, and Piga (2022) and Roveda, Castaman, Franceschi, Ghidoni and Pedrocchi (2020) respectively, Fig. 1), it can be written:

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q},\dot{\mathbf{q}}) = \boldsymbol{\mu} + \boldsymbol{\mu}_{ext}, \tag{1}$$

where $\mathbf{M}(\mathbf{q})$ is the inertia matrix of the robot, $\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})$ is the Coriolis vector, $\mathbf{q}$ is the joint position vector, $\boldsymbol{\mu}$ is the control torque vector, and $\boldsymbol{\mu}_{ext}$ is the external torque vector acting on the robot. The reference passive velocity field $V(\mathbf{q},\tau)$ (where τ is a parametrization variable) can be designed to guarantee the passivity of the controlled system, satisfying the following expression:

$$\int_0^t \boldsymbol{\mu}_{ext}^T \dot{\mathbf{q}} d\tau \geq W(\mathbf{q}(t),\dot{\mathbf{q}}(t)) - W(\mathbf{q}_0,\dot{\mathbf{q}}_0), \tag{2}$$

where $\boldsymbol{\mu}_{ext}^T\dot{\mathbf{q}}$ is the so-called supply rate, $W(\mathbf{q}(t),\dot{\mathbf{q}}(t))$ is a positive value function of the state s, and $W(\mathbf{q}_0,\dot{\mathbf{q}}_0) = c^2$, where $c \in \mathbb{R}$ satisfies (2) (depending on the initial conditions of the system $\mathbf{q}_0$ and $\dot{\mathbf{q}}_0$). To do so, the system dynamics in (1) is augmented with the dynamics of the evolution of the parametrization τ and with the dynamics of a fictitious energy storage element (*i.e.*, a fictitious flywheel):

$$\begin{bmatrix} \mathbf{M}(\mathbf{q}) & \mathbf{0}_{n\times 1} & \mathbf{0}_{n\times 1} \\ \mathbf{0}_{1\times n} & 1 & 0 \\ \mathbf{0}_{1\times n} & 0 & M_f \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}} \\ \ddot{\tau} \\ \ddot{q}_f \end{bmatrix} + \begin{bmatrix} \mathbf{C}(\mathbf{q},\dot{\mathbf{q}}) & \mathbf{0}_{n\times 1} & \mathbf{0}_{n\times 1} \\ \mathbf{0}_{1\times n} & 0 & 0 \\ \mathbf{0}_{1\times n} & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\tau} \\ \dot{q}_f \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mu} \\ \mu_\tau \\ \mu_f \end{bmatrix} + \begin{bmatrix} \boldsymbol{\mu}_{ext} \\ 0 \\ 0 \end{bmatrix}, \tag{3}$$

that is:

$$\mathbf{M}_{aug}(\mathbf{q}_{aug})\ddot{\mathbf{q}}_{aug} + \mathbf{C}_{aug}(\mathbf{q}_{aug},\dot{\mathbf{q}}_{aug}) = \boldsymbol{\mu}_{aug} + \boldsymbol{\mu}_{aug_{ext}}. \tag{4}$$

The reference passive velocity field $V(\mathbf{q},\tau)$ is then designed so that, when it is tracked, the kinetic energy of the augmented system in (4) will remain constant (*i.e.*, $\bar{E}(\mathbf{q}_{aug},\dot{\mathbf{q}}_{aug}) = \frac{1}{2}\dot{\mathbf{q}}_{aug}^T\mathbf{M}_{aug}\dot{\mathbf{q}}_{aug} = const$):

$$\mathbf{V}(\mathbf{q},\tau) = \mathbf{V}_{\parallel} + \mathbf{V}_{\perp} = \lambda_1(\mathbf{q},\tau)\begin{bmatrix} \frac{d\mathbf{q}^d(\tau)}{dt} \\ 1 \end{bmatrix} - \lambda_2(\mathbf{q},\tau)\begin{bmatrix} grad(U(\mathbf{E}(\mathbf{q},\tau))) \\ 0 \end{bmatrix}, \tag{5}$$

where $\mathbf{V}_{\parallel}$ is the component of the velocity field along the reference trajectory, and $\mathbf{V}_{\perp}$ is the component of the velocity field perpendicular to the reference trajectory. The ith element of $\mathbf{E}(\mathbf{q},\tau)$ is defined as it follows:

$$E_i(q_i,\tau) = e^{j\epsilon_i(q_i,\tau)}, \tag{6}$$

where $\epsilon_i(q_i,\tau) = (q^i - q_d^i(\tau))$. The potential field $U(\mathbf{E}(\mathbf{q},\tau))$ is defined as it follows:

$$U(\mathbf{E}(\mathbf{q},\tau)) = \sum_1^n h_i(1 - \cos(\epsilon_i(q_i,\tau))), \tag{7}$$

where $\mathbf{h}$ is a vector of weights. $\mathbf{q}^d$ is the reference joint position to be tracked. λ_1, λ_2 are weights defined as follows:

$$\lambda_1(\mathbf{q},\tau) = e^{-RU(\mathbf{E}(\mathbf{q},\tau))}, \tag{8}$$

$$\lambda_2(\mathbf{q},\tau) = 2 - e^{-RU(\mathbf{E}(\mathbf{q},\tau))}, \tag{9}$$

in which $R \geq$ is the so-called self-pacing parameter. The velocity field is, indeed, composed of two components, one towards the path and one along it: the velocity field prioritizes following one over the other depending on the distance of the actual joint position $\mathbf{q}$ w.r.t. the reference joint position $\mathbf{q}^d$. It can be seen that when the tracking error is large, the velocity field moves only towards the path while having a small tracking error the velocity field also moves along it. In fact, if $\mathbf{E}(\mathbf{q},\tau)$ is large, $\lambda_1(\boldsymbol{q},\tau)\rightarrow 0$ and $\lambda_2(\boldsymbol{q},\tau)\rightarrow 1$; conversely, if the error is close to zero $\mathbf{E}(\mathbf{q},\tau)\rightarrow 0$, $\lambda_1(\boldsymbol{q},\tau)\rightarrow 1$ and $\lambda_2(\boldsymbol{q},\tau)\rightarrow 1$.

Once the velocity field (5) is defined, it can be augmented with the one of the flywheel, obtaining $\mathbf{V}_{aug}(\mathbf{q},\tau) = [\mathbf{V}(\mathbf{q},\tau), V_f(\mathbf{q},\tau)]^T$. Therefore, by setting the value of the total kinetic energy of the augmented system in (4) $\bar{E} > 0$, the velocity field for the flywheel is computed as it follows:

$$V_f(\mathbf{q},\tau) = \sqrt{\frac{2}{M_f}\left(\bar{E} - \frac{1}{2}\mathbf{V}(\mathbf{q},\tau)^T\mathbf{M}_s(\mathbf{q})\mathbf{V}(\mathbf{q},\tau)\right)}, \tag{10}$$

where $\mathbf{M}_s(\mathbf{q})$ is the mass matrix for the suspended system (*i.e.*, not considering the flywheel dynamics in (4)):

$$\mathbf{M}_s(\mathbf{q}) = \begin{bmatrix} \mathbf{M}(\mathbf{q}) & \mathbf{0}_{n\times 1} \\ \mathbf{0}_{1\times n} & 1. \end{bmatrix} \tag{11}$$

$\bar{E}$ should be selected high enough to allow the above equation to have a real solution, meaning that the energy stored in the fictitious flywheel should be high enough to cover the energy demand at any point of the reference path.

Once the augment velocity field $\mathbf{V}_{aug}(\mathbf{q},\tau)$ is defined, the control action is defined on the basis of the following quantities:

$$\mathbf{p}_{aug}(\mathbf{q}_{aug},\dot{\mathbf{q}}_{aug}) = \mathbf{M}_{aug}(\mathbf{q}_{aug})\dot{\mathbf{q}}_{aug}, \tag{12}$$

which is the real momentum of the augmented system;

$$\mathbf{P}_{aug}(\mathbf{q}_{aug}) = \mathbf{M}_{aug}(\mathbf{q}_{aug})\mathbf{V}_{aug}(\mathbf{q}_{aug}), \tag{13}$$

which is the desired momentum of the augmented system;

$$\mathbf{w}_{aug}(\mathbf{q}_{aug},\dot{\mathbf{q}}_{aug}) = \mathbf{M}_{aug}(\mathbf{q}_{aug})\dot{\mathbf{V}}_{aug}(\mathbf{q}_{aug}) + \mathbf{C}_{aug}(\mathbf{q}_{aug},\dot{\mathbf{q}}_{aug})\mathbf{V}_{aug}(\mathbf{q}_{aug}), \tag{14}$$

which is the covariant derivative of the desired momentum. The control law is then defined as follows:

$$\boldsymbol{\mu}_{aug} = \boldsymbol{\mu}_f + \boldsymbol{\mu}_c, \tag{15}$$

where:

$$\boldsymbol{\mu}_f = \gamma(\mathbf{P}_{aug}\mathbf{p}_{aug}^T - \mathbf{p}_{aug}\mathbf{P}_{aug}^T)\dot{\mathbf{q}}_{aug}, \tag{16}$$

$$\boldsymbol{\mu}_c = \frac{1}{2E}(\mathbf{w}_{aug}\mathbf{P}_{aug}^T - \mathbf{P}_{aug}\mathbf{w}_{aug}^T)\dot{\mathbf{q}}_{aug}. \tag{17}$$

$\boldsymbol{\mu}_f$ can be seen as a feedback control since it depends on the relation between the desired and actual momentum. γ is thereby the feedback gain, and it can be used as a tuning parameter when implementing the control law. $\boldsymbol{\mu}_c$ can be seen as a feed-forward action.

In this work, the PVFC is used to track the reference joint position and velocity coming from the intermediate control loop, that is the Cartesian impedance control (Fig. 1). In order to ensure good tracking performance in time, the self-pacing parameter R is set to zero. The designed velocity field in (5) becomes:

$$\mathbf{V}(\mathbf{q},\tau) = \begin{bmatrix} \frac{d\mathbf{q}^{\mathbf{d}}(\tau)}{dt} \\ 1 \end{bmatrix} - \begin{bmatrix} grad(U(\mathbf{E}(\mathbf{q},\tau))) \\ 0 \end{bmatrix}. \tag{18}$$

In addition, a feedback control loop is implemented on the total kinetic energy of the system $\bar{E}(\mathbf{q}_{aug},\dot{\mathbf{q}}_{aug})$. Thus, a PI feedback controller is designed to act on the total energy error $e_{\bar{E}}$, adapting the flywheel control action μ_f.

Remark 2. It is important to highlight that the PVFC is used instead of a standard position/velocity inner control loop to design a passive low-level tracking controller. In fact, impedance controllers commanding

the motor torques might suffer from tracking issues due to the intrinsic compliant behavior of this controller.

Remark 3. Ideally, by properly compensating the joint friction effects, the parametrization variable τ is always equal to the time t.

2.2. Cartesian impedance control

As described in Caccavale et al. (1999), a Cartesian impedance controller can be designed to provide the robot with a compliant behavior, which is the aim of the intermediate loop (Fig. 1). On the basis of the interaction force acting on the manipulator, impedance control allows calculating the robot accelerations $\ddot{\mathbf{x}}_{imp} = [\ddot{\mathbf{p}}; \ddot{\boldsymbol{\varphi}}_{cd}]$ (where $\ddot{\mathbf{p}}$ are related to the transnational degrees of freedom – DoFs –, and $\ddot{\boldsymbol{\varphi}}_{cd}$ are related to the rotational DoFs described by the intrinsic Euler angles representation):

$$\begin{aligned} \ddot{\mathbf{p}} &= \mathbf{M}_t^{-1}\left(-\mathbf{D}_t \Delta\dot{\mathbf{p}} - \mathbf{K}_t \Delta\mathbf{p} - \mathbf{f}_t\right), \\ \ddot{\boldsymbol{\varphi}}_{cd} &= \mathbf{M}_\varphi^{-1}\left(-\mathbf{D}_\varphi \boldsymbol{d}\dot{\boldsymbol{\varphi}}_{cd} - \mathbf{K}_\varphi \boldsymbol{\varphi}_{cd} + \mathbf{S}_\omega^T(\boldsymbol{\varphi}_{cd})\,\boldsymbol{\mu}_\varphi\right). \end{aligned} \tag{19}$$

Considering the transnational part of the impedance control, $\mathbf{M}_t$ is the target mass matrix, $\mathbf{D}_t$ is the target damping matrix, $\mathbf{K}_t$ is the target stiffness matrix, $\mathbf{f}_t$ is the external forces vector. $\mathbf{p}$ is the actual Cartesian positions vector, while $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^d$ and $\Delta\dot{\mathbf{p}} = \dot{\mathbf{p}} - \dot{\mathbf{p}}^d$, where $\mathbf{p}^d$ is the target positions vector and $\dot{\mathbf{p}}^d$ is the target velocity vector. Considering the rotational part of the impedance control, $\mathbf{M}_\varphi$ is the target inertia matrix, $\mathbf{D}_\varphi$ is the target damping matrix, $\mathbf{K}_\varphi$ is the target stiffness matrix. $\boldsymbol{\varphi}_{cd}$ is the set of Euler angles extracted from $\mathbf{R}_c^d = \mathbf{R}_d^T\mathbf{R}_c$, describing the mutual orientation between the compliant frame $\mathbf{R}_c$ (coincident with the robot end-effector reference frame) and the target frame $\mathbf{R}_d$. $\boldsymbol{\mu}_\varphi$ is the external torques vector referred to the target frame. Matrix $\mathbf{S}_\omega(\boldsymbol{\varphi}_{cd})$ defines the transformation from Euler angles derivatives to angular velocities $\boldsymbol{\omega} = \mathbf{S}_\omega(\boldsymbol{\varphi}_{cd})\boldsymbol{d}\dot{\boldsymbol{\varphi}}_{cd}$ (Sciavicco & Siciliano, 2012). The six DoFs impedance control results, therefore, in:

$$\mathbf{M}\,\ddot{\mathbf{x}}_{Ci} + \mathbf{D}\,\Delta\dot{\mathbf{x}}_{Ci} + \mathbf{K}\,\Delta\mathbf{x}_{Ci} = \mathbf{h}_{ext}, \tag{20}$$

where $\mathbf{M}$, $\mathbf{D}$, $\mathbf{K}$ are the impedance diagonal matrices composed by both the transnational and rotational parts, $\Delta\mathbf{x}_{Ci} = \mathbf{x} - \mathbf{x}^d = [\Delta\mathbf{p}; \boldsymbol{\varphi}_{cd}] = [\Delta\mathbf{x}_t; \Delta\mathbf{x}_r]$ (where $\mathbf{x}^d = [\mathbf{x}_t^d, \mathbf{x}_r^d]$ is the six DoFs position reference for the impedance controller, with transnational $\mathbf{x}_t^d$ and rotation $\mathbf{x}_r^d$ components), $\Delta\dot{\mathbf{x}}_{Ci} = \dot{\mathbf{x}} - \dot{\mathbf{x}}^d$, and $\mathbf{h}_{ext} = [\mathbf{f}_t; \mathbf{S}_\omega^T(\boldsymbol{\varphi}_{cd})\,\boldsymbol{\mu}_\varphi]$. It has to be underlined that the damping matrix can be computed as it follows: $\mathbf{D} = 2\zeta\sqrt{\mathbf{KM}}$, where $\zeta = diag([\zeta_t, \zeta_r])$ (where ζ_t and ζ_r are related to the transnational and rotational part of the Cartesian impedance control, respectively) is the damping ratio diagonal matrix.

On the basis of the initial conditions on the impedance control position ($\mathbf{x}_{Ci}(t = 0) = \mathbf{x}_{Ci}^0$) and velocity $\dot{\mathbf{x}}_{Ci}(t = 0) = \dot{\mathbf{x}}_{Ci}^0$), it is possible to compute the impedance control acceleration $\ddot{\mathbf{x}}_{Ci}$ from (20). Such acceleration can be integrated in order to compute the Cartesian position $\mathbf{x}_{Ci}$ and velocity (considering angular velocities for the rotational components) $\dot{\mathbf{x}}_{Ci}$ to be used as a reference to the inner position controller. The Cartesian position $\mathbf{x}_{Ci}$ and velocity $\dot{\mathbf{x}}_{Ci}$ are then used to compute the joint reference signals to be provided to the PVFC. The joint reference velocity vector $\dot{\mathbf{q}}^d$ is computed as it follows (Siciliano, Sciavicco, Villani, & Oriolo, 2010):

$$\dot{\mathbf{q}}^d = \mathbf{J}^{\#}(\mathbf{q})\dot{\mathbf{x}}_{Ci}, \tag{21}$$

where $^{\#}$ denotes the pseudoinverse. Form (21), the reference joint position $\mathbf{q}^d$ can be computed (*i.e.*, integrating the reference joint velocity $\dot{\mathbf{q}}^d$), to be provided to the PVFC.

2.3. Optimal linear quadratic regulator

2.3.1. Controller design

The outer control layer is designed to make use of the LQR formulation (Fig. 1). Such a control loop aims to detect the intention of motion of the human, making the robot react to minimize the interaction force. To do that, the Cartesian impedance control setpoint is deformed. The LQR is applied to the translational degrees of freedom (DoFs) to enable human–robot collaboration (Roveda, Maskani, et al., 2020). In fact, due to the higher complexity, rotational DoFs can be managed (if needed) with different strategies to perform the target task (Roveda, Haghshenas, et al., 2018). In order to design the proposed LQR, the human's arm dynamics has to be modeled, coupling it with the robot-controlled dynamics in (20). Based on Artemiadis et al. (2010), the human's arm dynamics is simply modeled as an elastic element (*i.e.*, as a spring, neglecting the damping term for the sake of stability). Considering a single DoF formulation, the force applied by the human $F_{h,i}$ can be written as it follows:

$$F_{h,i} = K_{h,i}(x_{h,i}^0 - x_{h,i}), \tag{22}$$

where $K_{h,i}$ is the equivalent stiffness of the human arm, $x_{h,i}$ is the current position of the human arm, and $x_{h,i}^0$ is the equilibrium position of the human arm. $x_{h,i} = x_{t,i}$ can be assumed since the human is considered to interact with the robot at its end-effector. By considering the one DoF formulation of the translation part of (20) and combining it with (22) (having $f_{t,i} = F_{h,i}$), the coupled human–robot interaction dynamics is described by:

$$M_{t,i}\ddot{x}_{t,i} + D_{t,i}\dot{x}_{t,i} + (K_{t,i} + K_{h,i})x_{t,i} = K_{t,i}x_{t,i}^d + K_{h,i}x_{h,i}^0. \tag{23}$$

Similarly to Roveda et al. (2015) (in which an optimal controller has been designed for robot-environment interaction tasks), it is possible to make use of (22) and (23) to define the following state-space model for the design of the LQR:

$$\begin{bmatrix} \ddot{x}_{t,i} \\ \dot{x}_{t,i} \\ \dot{e}_{F_{h,i}} \end{bmatrix} = \begin{bmatrix} -\frac{D_{t,i}}{M_{t,i}} & -\frac{(K_{t,i}+K_{h,i})}{M_{t,i}} & 0 \\ 1 & 0 & 0 \\ 0 & K_{h,i} & 0 \end{bmatrix} \begin{bmatrix} \ddot{x}_{t,i} \\ \dot{x}_{t,i} \\ e_{F_{h,i}} \end{bmatrix} + \begin{bmatrix} \frac{K_{t,i}}{M_{t,i}} \\ 0 \\ 0 \end{bmatrix} \dot{x}_{t,i}^d, \tag{24}$$

where $e_{F_{h,i}} = F_i^d - F_{h,i} = F_i^d - K_{h,i}(x_{h,i}^0 - x_{h,i})$, and F_i^d is the reference force. The human's setpoint $x_{h,i}^0$ is considered to be updated with a low-frequency w.r.t. the bandwidth of the proposed controller (*i.e.*, the human defines where to go, keeping such reference position fixed for a certain time Fligge, McIntyre, & van der Smagt, 2012), so that its variation is negligible for the controller design in (24) (*i.e.*, $\dot{x}_{h,i}^0 = 0$). The same assumption can be made on the variation of the human's arm stiffness (Kronander & Billard, 2012), *i.e.*, neglecting the variation of $K_{h,i}$ in (24) (*i.e.*, $\dot{K}_{h,i} = 0$). $F_i^d = 0\ N$ is imposed since the aim of the controller is to minimize the interaction force between the human and the robot to achieve a smooth and effortless collaboration. (24) can be rewritten as:

$$\dot{\boldsymbol{\eta}}_i = \mathbf{A}_i\boldsymbol{\eta}_i + \mathbf{b}_i\dot{x}_{t,i}^d. \tag{25}$$

Despite the variation of the human's arm stiffness $\dot{K}_{h,i}$ is negligible, for the implementation of a robust controller w.r.t. the human's arm stiffness $K_{h,i}$ (therefore considering the possibility to have the human adapting his/her arm's stiffness), it is possible to include such uncertainties into the proposed modeling (25):

$$\dot{\boldsymbol{\eta}}_i = \mathbf{A}_i(U_{K_{h,i}})\boldsymbol{\eta}_i + \mathbf{b}_i\dot{x}_{t,i}^d, \tag{26}$$

where $U_{K_{h,i}}$ models the uncertainties bounds of the human's arm stiffness $K_{h,i}$, and:

$$\mathbf{A}_i(U_{K_{h,i}}) = \begin{bmatrix} -\frac{D_i}{M_i} & -\frac{(K_i+K_{h,i}+U_{K_{h,i}})}{M_i} & 0 \\ 1 & 0 & 0 \\ 0 & K_{h,i} + U_{K_{h,i}} & 0 \end{bmatrix}. \tag{27}$$

The controller performance index needed to design the control is:

$$J_{cost,i} := \int_0^\infty (\boldsymbol{\eta}_i^T\mathbf{Q}_{opt,i}\boldsymbol{\eta}_i + (\dot{x}_{t,i}^d)^2)dt. \tag{28}$$

The matrix $\mathbf{Q}_{opt,i} := diag(q_{1,i}, q_{2,i}, q_{3,i})$ is the LQR gain matrix, where $(q_{1,i}, q_{2,i}, q_{3,i}) \geq 0$ can be tuned in order to modulate the control gains. More specifically, $q_{1,i}$ is the gain related to the acceleration $\ddot{x}_{t,i}$, $q_{2,i}$ is the gain related to the velocity $\dot{x}_{t,i}$, and $q_{3,i}$ is the gain related to the force error $e_{F_{h,i}}$. In particular, $q_{1,i}$ and $q_{2,i}$ are imposed equal to 0, while $q_{3,i}$ is used to manage the LQR output (*i.e.*, the deformation of the Cartesian impedance control setpoint). Indeed, the control action is optimized to minimize the cost function in (28):

$$J_i^*(\boldsymbol{\eta}_i) := \min_{\dot{x}^d_{opt,t,i}} \{J_{cost,i}(\boldsymbol{\eta}_i)\}, \tag{29}$$

which corresponds to solving the related Riccati matrix equation:

$$\begin{aligned}\mathbf{0} &= -\mathbf{S}_i\mathbf{b}_i\mathbf{b}_i^T\mathbf{S}_i \\ &+ \mathbf{S}_i\mathbf{A}_i(U_{K_{h,i}}) + \mathbf{F}_{u,i} + \beta_{u,i}^2\mathbf{I} \\ &+ \mathbf{A}_i^T(U_{K_{h,i}})\mathbf{S}_i + \mathbf{Q}_{opt,i},\end{aligned} \tag{30}$$

where the solution is the symmetric 3×3 constant matrix $\mathbf{S}_i$, $\beta_{u,i}$ is a design parameter, and $\mathbf{F}_{u,i}$ includes the modeling uncertainties. In particular, $\boldsymbol{F}_{u,i}$ is defined as it follows:

$$\mathbf{F}_{u,i} = \boldsymbol{\Delta}\mathbf{A}_i^T(\mathbf{b}_i^{\#})^T\mathbf{b}_i^{\#}\boldsymbol{\Delta}\mathbf{A}_i, \tag{31}$$

where:

$$\boldsymbol{\Delta}\mathbf{A}_i = \begin{bmatrix} 0 & -\frac{U_{K_{h,i}}}{M_i} & 0 \\ 0 & 0 & 0 \\ 0 & U_{K_{h,i}} & 0 \end{bmatrix} \tag{32}$$

The impedance reference velocity can, therefore, be computed:

$$\dot{x}^d_{t,i} = -\mathbf{b}_i^T\mathbf{S}_i\boldsymbol{\eta}_i. \tag{33}$$

The solution matrix $\mathbf{S}_i$ can be analytically computed and it is, intrinsically, dependent on the control parameters $(D_{t,i}, K_{t,i}, q_{3,i})$ and on the uncertainties related to the human's arm stiffness $K_{h,i}$. Indeed, by computing the analytical LQR solution, it can be easily implemented and calculated online based on the imposed control/modeling parameters.

The analytic solution of (30) is presented below. By defining $\omega_i = \sqrt{\frac{K_{t,i}}{M_{t,i}}}$ and $K^{tot}_{t,i} = K_{t,i} + K_{h,i}$, the following six equations can be written:

$$\begin{cases} 0 = & -S_i(1,1)^2\omega_i^4 - 2S_i(1,1)\frac{D_{t,i}}{M_{t,i}} + 2S_i(1,2) \\ & + q_{1,i} + \beta_{u,i}^2, \\ 0 = & -S_i(1,1)S_i(1,2)\omega_i^4 + S_i(2,2) \\ & + S_i(1,3)(K_{h,i} + U_{K_{h,i}}) - S_i(1,2)\frac{D_{t,i}}{M_{t,i}} \\ & - S_i(1,1)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}}, \\ 0 = & -S_i(1,1)S_i(1,3)\omega_i^4 + S_i(2,3) - S_i(1,3)\frac{D_{t,i}}{M_{t,i}}, \\ 0 = & -S_i(1,2)^2\omega_i^4 - 2S_i(1,2)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}} \\ & + 2S_i(2,3)(K_{h,i} + U_{K_{h,i}}) + q_{2,i} + \frac{U^2_{K_{h,i}}}{M^2_{t,i}\omega_i^4} + \beta_{u,i}^2, \\ 0 = & -S_i(1,2)S_i(1,3)\omega_i^4 + S_i(3,3)(K_{h,i} + U_{K_{h,i}}) \\ & - S_i(1,3)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}}, \\ 0 = & -S_i(1,3)^2\omega_i^2 + q_{3,i} + \beta_{u,i}^2. \end{cases} \tag{34}$$

From the last equation of (34), $S_i(1,3)$ can be obtained:

$$S_i(1,3) = -\frac{\sqrt{q_{3,i} + \beta_{u,i}^2}}{\omega_i^2}. \tag{35}$$

Substituting (35) into (34):

$$\begin{cases} 0 = & -S_i(1,1)^2\omega_i^4 - 2S_i(1,1)\frac{D_{t,i}}{M_{t,i}} + 2S_i(1,2) \\ & + q_{1,i} + \beta_{u,i}^2, \\ 0 = & -S_i(1,1)S_i(1,2)\omega_i^4 + S_i(2,2) \\ & - \frac{\sqrt{q_{3,i}+\beta_{u,i}^2}}{\omega_i^2}(K_{h,i} + U_{K_{h,i}}) - S_i(1,2)\frac{D_{t,i}}{M_{t,i}} \\ & - S_i(1,1)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}}, \\ 0 = & +S_i(1,1)\sqrt{q_{3,1} + \beta_{u,i}^2}\omega_i^2 + S_i(2,3) \\ & + \frac{\sqrt{q_{3,i}+\beta_{u,i}^2}}{\omega_i^2}\frac{D_{t,i}}{M_{t,i}}, \\ 0 = & -S_i(1,2)^2\omega_i^4 - 2S_i(1,2)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}} \\ & + 2S_i(2,3)(K_{h,i} + U_{K_{h,i}}) + q_{2,i} + \frac{U^2_{K_{h,i}}}{M^2_{t,i}\omega_i^4} + \beta_{u,i}^2, \\ 0 = & +S_i(1,2)\sqrt{q_{3,1} + \beta_{u,i}^2}\omega_i^2 \\ & + S_i(3,3)(K_{h,i} + U_{K_{h,i}}) + \frac{\sqrt{q_{3,i}+\beta_{u,i}^2}}{\omega_i^2}\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}}. \end{cases} \tag{36}$$

From the second, the third, and the fifth equations of (36) it is possible to obtain:

$$\begin{aligned} S_i(2,2) = & + S_i(1,1)S_i(1,2)\omega_i^4 + S_i(1,2)\frac{D_{t,i}}{M_{t,i}} \\ & + S_i(1,1)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}} \\ & + \frac{\sqrt{q_{3,i} + \beta_{u,i}^2}}{\omega_i^2}(K_{h,i} + U_{K_{i,h}}), \end{aligned} \tag{37}$$

$$\begin{aligned} S_i(2,3) = & -\frac{D_{t,i}}{M_{t,i}}\frac{\sqrt{q_{3,i} + \beta_{u,i}^2}}{\omega_i^2} \\ & - S_i(1,1)\sqrt{q_{3,1} + \beta_{u,i}^2}, \end{aligned} \tag{38}$$

$$\begin{aligned} S_i(3,3) = & -S_i(1,2)\frac{\sqrt{q_{3,i} + \beta_{u,i}^2}\omega_i^2}{K_{h,i} + U_{K_{h,i}}} \\ & - \frac{\sqrt{q_{3,i} + \beta_{u,i}^2}}{\omega_i^2 M_{t,i}}\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{K_{h,i} + U_{K_{h,i}}}. \end{aligned} \tag{39}$$

Substituting (39) into (36):

$$\begin{cases} 0 = & -S_i(1,1)^2\omega_i^4 - 2S_i(1,1)\frac{D_{t,i}}{M_{t,i}} + 2S_i(1,2) \\ & + q_{1,i} + \beta_{u,i}^2, \\ 0 = & -S_i(1,2)^2\omega_i^4 - 2S_i(1,2)\frac{K^{tot}_{t,i} + U_{K_{h,i}}}{M_{t,i}} \\ & - 2\frac{D_{t,i}}{M_{t,i}}\frac{\sqrt{q_{3,i}+\beta_{u,i}^2}}{\omega_i^2}(K_{h,i} + U_{K_{h,i}}) \\ & - S_i(1,1)\sqrt{q_{3,1} + \beta_{u,i}^2}(K_{h,i} + U_{K_{h,i}}) \\ & + q_{2,i} + \frac{U^2_{K_{h,i}}}{M_i^2\omega_i^4} + \beta_{u,i}^2. \end{cases} \tag{40}$$

$S_i(1,2)$ can be then calculated from the first equation of (40):

$$S_i(1,2) = S_i(1,1)^2\frac{\omega_i^4}{2} + S(1,1)\frac{D_{t,i}}{M_{t,i}} - \frac{q_{1,i}}{2} - \frac{\beta_{u,i}^2}{2}. \tag{41}$$

Finally, by substituting (41) into (40), the following fourth order equation can be found:

$$0 = F_{1,i}S_i(1,1)^4 + F_{2,i}S_i(1,1)^3 + F_{3,i}S_i(1,1)^2 + F_{4,i}S_i(1,1) + F_{5,i}, \tag{42}$$

where:

$$\begin{cases} F_{1,i} = & -\frac{\omega_i^{12}}{4}, \\ F_{2,i} = & -\omega_i^8 \frac{D_{t,i}}{M_{t,i}}, \\ F_{3,i} = & -(\frac{D_{t,i}}{M_{t,i}})^2 \omega_i^4 + \frac{\omega_i^8}{2}(q_{1,i} + \beta_{u,i}^2) \\ & - \omega_i^4 \frac{K_{t,i}^{tot} + U_{K_{h,i}}}{M_{t,i}}, \\ F_{4,i} = & + \frac{D_{t,i}}{M_{t,i}}(q_{1,i} + \beta_{u,i}^2)\omega_i^4 - 2\frac{D_{t,i}}{M_{t,i}^2}(K_{t,i}^{tot} + U_{K_{h,i}}) \\ & - \sqrt{q_{3,1} + \beta_{u,i}^2}(K_{h,i} + U_{K_{h,i}}), \\ F_{5,1} = & -\frac{q_{1,i}^2}{4}\omega_i^4 - \frac{\beta_{u,i}^2}{4}\omega_i^4 - \frac{q_{1,i}\beta_{u,i}^2}{2}\omega_i^4 \\ & + (q_{1,i} + \beta_{u,i}^2)\frac{K_{t,i}^{tot} + U_{K_{h,i}}}{M_{t,i}} \\ & - 2\frac{D_{t,i}}{M_{t,i}}\frac{\sqrt{q_{3,i} + \beta_{u,i}^2}}{\omega_i^2}(K_{h,i} + U_{K_{h,i}}) \\ & + q_{2,1} + \frac{U_{K_{h,i}}^2}{M_{t,i}^2 \omega_i^4} + \beta_{u,i}^2. \end{cases} \tag{43}$$

To solve this fourth-order equation many methods exist, both analytical and numerical; below an analytical is proposed. Let us define:

$$\begin{cases} p_{1,i} = & \frac{8F_{1,i}F_{3,i} - 3F_{2,i}^2}{8F_{1,i}^2}, \\ q_{1,i} = & \frac{F_{2,i}^3 - 4F_{1,i}F_{2,i}F_{3,i} + 8F_{1,i}^2 F_{4,i}}{8F_{1,i}^3}, \\ \Delta_{0,i} = & F_{3,i}^2 - 3F_{2,i}F_{4,i} + 12F_{1,i}F_{5,i}, \\ \Delta_{1,i} = & 2F_{3,i}^3 - 9F_{2,i}F_{3,i}F_{4,i} + 27F_{2,i}^2 F_{5,i} \\ & + 27F_{1,i}F_{4,i}^2 - 72F_{1,i}F_{3,i}F_{5,i}, \\ V_i = & \sqrt[\frac{1}{3}]{\Delta_{1,i} + \frac{\sqrt{\Delta_{1,i}^2 - 4\Delta_{0,i}^3}}{2}}, \\ N_i = & \frac{\sqrt{-\frac{2}{3}p_{1,i} + \frac{V_i + \frac{\Delta_{0,i}}{V_i}}{3F_{1,i}}}}{2}. \end{cases} \tag{44}$$

The solutions are:

$$\begin{cases} S_i(1,1)^{\#1} = -\frac{F_{2,i}}{4F_{1,i}} - N_i + \frac{\sqrt{-4N_i^2 - 2p_{1,i} + \frac{q_{1,i}}{N_i}}}{2}, \\ S_i(1,1)^{\#2} = -\frac{F_{2,i}}{4F_{1,i}} - N_i - \frac{\sqrt{-4N_i^2 - 2p_{1,i} + \frac{q_{1,i}}{N_i}}}{2}, \\ S_i(1,1)^{\#3} = -\frac{F_{2,i}}{4F_{1,i}} + N_i + \frac{\sqrt{-4N_i^2 - 2p_{1,i} - \frac{q_{1,i}}{N_i}}}{2}, \\ S_i(1,1)^{\#4} = -\frac{F_{2,i}}{4F_{1,i}} + N_i - \frac{\sqrt{-4N_i^2 - 2p_{1,i} - \frac{q_{1,i}}{N_i}}}{2}. \end{cases} \tag{45}$$

The selected $S_i(1,1)$ will be the real positive solution, and it will be used in (37), (38), (39), and (41) to compute the remaining elements of $\mathbf{S}_i$.

Once the control gains are calculated, the Cartesian impedance control setpoint $x_{t_i}^d$ can be computed by integrating $\dot{x}_{t,i}^d$ in (33) at each control step.

Remark 4. Çimen (2008) provides details on how to modify the proposed controller to include given constraints (*e.g.*, on the task space).

2.3.2. Stability analysis

To prove the stability of the LQR control, it is possible to refer to an analogous LTI system without the final penalty term and to transform it into an infinite-horizon linear quadratic optimal control problem (same procedure applied to the developed control). Note that this is possible only under the assumption of the controllability of the LTI system (Fisher & Bhattacharya, 2009).

Let us consider the LTI system defined in (26) and the matrices defined in (28):

- the system is controllable because the reachability matrix $\mathbf{RM}_i = [\mathbf{b}_i \;\; \mathbf{A}_i(U_{K_{h,i}})\mathbf{b}_i \;\; \mathbf{A}_i(U_{K_{h,i}})^2\mathbf{b}_i]$ is not rank deficient;
- the matrix $\mathbf{Q}_{opt,i}$ is semi positive definite, and the matrix $[\mathbf{R}_{opt,i}] = [I]$ is positive definite.

Indeed, the controlled system is asymptotically stable.

2.4. Global stability of the control schema

Having that the LQR controller is asymptotically stable, the stability of the impedance control is guaranteed (Colgate & Hogan, 1989). In fact, the impedance controller mass, stiffness, and damping parameters are constant (*i.e.*, the preference-based optimization adapts the control gains before the controller is activated). The impedance controller setpoint is provided by the LQR controller, being asymptotically stable and, therefore, not affecting the stability of the impedance control loop.

The exponential stability of the PVFC is provided in Li and Horowitz (2001b), even in the presence of external forces. Indeed, having as an input an asymptotically stable setpoint (provided by the impedance controller), the stability of this control loop is guaranteed.

3. Preference-based optimization

The PBO algorithm employed in this paper is based on the methodology developed in Bemporad and Piga (2021) by some of the authors, where the GLISp algorithm is introduced. Such an approach has been used to optimize the gains of the proposed controller, especially considering the Cartesian impedance control stiffness $K_{t,i}$ and damping $\zeta_{t,i}$, and the LQR gain $q_{3,i}$. In such a way, a human-centric tuning procedure is implemented, capable to catch the (even intrinsic) human's needs. In the following, the PBO algorithm is described.

3.1. Building a surrogate function from preferences

The first step of the GLISp algorithm is to build a surrogate function describing the observed preferences.

Formally, given two possible sets of parameters θ_1 and θ_2, the *preference function* $\pi : \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_\theta} \to \{-1, 0, 1\}$ is defined as:

$$\pi(\theta_1, \theta_2) = \begin{cases} -1 & \text{if } \theta_1 \text{ “better” than } \theta_2 \\ 0 & \text{if } \theta_1 \text{ “as good as” } \theta_2 \\ 1 & \text{if } \theta_2 \text{ “better” than } \theta_1. \end{cases} \tag{46}$$

Assuming that $N \geq 2$ samples $\{\theta_1 \; \dots \; \theta_N\}$ of the decision vector are generated, with $\theta_i, \theta_j \in \mathbb{R}^{n_\theta}$ such that $\theta_i \neq \theta_j$, $\forall i \neq j$, with $i, j = 1, \dots, N$. For each of these parameters, an experiment is performed and the user provides a *preference vector* $\mathbf{B} = [b_1 \; \dots \; b_M]^T \in \{-1, 0, 1\}^M$ with:

$$b_h = \pi(\theta_{i(h)}, \theta_{j(h)}), \tag{47}$$

where M is the number of expressed preferences, $h \in \{1, \dots, M\}$, $i(h), j(h) \in \{1, \dots, N\}$, and $i(h) \neq j(h)$. It has to be noted that the element b_h of the vector the $\mathbf{B}$ represents the preference expressed by the user between the experimental performance achieved with the parameters $\theta_{i(h)}$ and $\theta_{j(h)}$.

The observed preferences are then used to learn a surrogate function $\hat{J} : \mathbb{R}^{n_\theta} \to \mathbb{R}$ of an (unknown) underlying performance index J. The surrogate $\hat{J}$ is parametrized as the following linear combination of Radial Basis Functions (RBFs):

$$\hat{J}(\theta) = \sum_{k=1}^{N} \beta_k \phi(\gamma d(\theta, \theta_i)), \tag{48}$$

where $d : \mathbb{R}^{n_\theta} \times \mathbb{R}^{n_\theta} \to \mathbb{R}$ is the squared Euclidean distance:

$$d(\boldsymbol{\theta}, \boldsymbol{\theta}_i) = \|\boldsymbol{\theta} - \boldsymbol{\theta}_i\|_2^2, \tag{49}$$

$\gamma > 0$ is a scalar parameter, $\phi : \mathbb{R} \to \mathbb{R}$ is an RBF, and $\boldsymbol{\beta} = [\beta_1 \ \dots \ \beta_N]^T$ are the unknown coefficients to be computed based on the available user's preferences. Examples of RBFs are $\phi(\gamma d) = \frac{1}{1+(\gamma d)^2}$ (*inverse quadratic*) and $\phi(\gamma d) = e^{-(\gamma d)^2}$ (*Gaussian*) (see Bemporad, 2020, for more definitions).

According to the preference relation (46), the surrogate $\hat{J}$ has to satisfy the following constraints:

$$\begin{array}{ll} \hat{J}(\boldsymbol{\theta}_{i(h)}) \le \hat{J}(\boldsymbol{\theta}_{j(h)}) - \sigma + \varepsilon_h & \text{if } \pi(\boldsymbol{\theta}_{i(h)}, \boldsymbol{\theta}_{j(h)}) = -1 \\ \hat{J}(\boldsymbol{\theta}_{i(h)}) \ge \hat{J}(\boldsymbol{\theta}_{j(h)}) + \sigma - \varepsilon_h & \text{if } \pi(\boldsymbol{\theta}_{i(h)}, \boldsymbol{\theta}_{j(h)}) = 1 \\ |\hat{J}(\boldsymbol{\theta}_{i(h)}) - \hat{J}(\boldsymbol{\theta}_{j(h)})| \le \sigma + \varepsilon_h & \text{if } \pi(\boldsymbol{\theta}_{i(h)}, \boldsymbol{\theta}_{j(h)}) = 0 \end{array} \tag{50}$$

for all $h = 1, \dots, M$, where $\sigma > 0$ is a given tolerance and ε_h is a positive slack variable that is used to relax the preference constraints. Constraints of infeasibility might be due to an inappropriate selection of the RBF (namely, poor flexibility in the parametric description of the surrogate $\hat{J}$) and/or inconsistent assessments done by the user.

Based on the above preference constraints, the coefficient vector $\boldsymbol{\beta}$ describing the surrogate $\hat{J}$ is obtained by solving the Quadratic Programming (QP) problem:

$$\begin{aligned} \min_{\beta,\varepsilon} \quad & \sum_{h=1}^{M} \varepsilon_h + \frac{\lambda}{2} \sum_{k=1}^{N} \beta_k^2 \\ \text{s.t.} \quad & \sum_{k=1}^{N} (\phi(\gamma d(\boldsymbol{\theta}_{i(h)}, \boldsymbol{\theta}_k))) - \phi(\gamma d(\boldsymbol{\theta}_{j(h)}, \boldsymbol{\theta}_k))\beta_k \\ & \qquad \le -\sigma + \varepsilon_h, \qquad \forall h : \ b_h = -1 \\ & \sum_{k=1}^{N} (\phi(\gamma d(\boldsymbol{\theta}_{i(h)}, \boldsymbol{\theta}_k))) - \phi(\gamma d(\boldsymbol{\theta}_{j(h)}, \boldsymbol{\theta}_k))\beta_k \\ & \qquad \ge \sigma - \varepsilon_h, \qquad \forall h : \ b_h = 1 \\ & \left| \sum_{k=1}^{N} (\phi(\gamma d(\boldsymbol{\theta}_{i(h)}, \boldsymbol{\theta}_k))) - \phi(\gamma d(\boldsymbol{\theta}_{j(h)}, \boldsymbol{\theta}_k))\beta_k \right| \\ & \qquad \le \sigma + \varepsilon_h, \qquad \forall h : \ b_h = 0 \\ & h = 1, \dots, M. \end{aligned} \tag{51}$$

The scalar $\lambda > 0$ in the cost function (51) is a regularization parameter that guarantees uniqueness in the solution of the QP problem.

3.2. Acquisition function

Once a surrogate $\hat{J}$ is estimated, this function can be minimized in order to find the optimal parameter vector $\boldsymbol{\theta}$.

More in detail, the following steps can be followed:
(i) generate a new sample by pure minimization of the estimated surrogate function $\hat{J}$, *i.e.*,

$$\boldsymbol{\theta}_{N+1} = \arg\min \hat{J}(\boldsymbol{\theta}) \ \text{ s.t. } \ \boldsymbol{\theta} \in \Theta;$$

(ii) ask the user to express a preference $\pi(\boldsymbol{\theta}_{N+1}, \boldsymbol{\theta}_N^\star)$, where $\boldsymbol{\theta}_N^\star \in \mathbb{R}^{n_\theta}$ is the vector of parameters with the best experimental output (judged by the user) found so far;
(iii) update the estimate of $\hat{J}$ through (51);
(iv) iterate over N.

Such a procedure, which only *exploits* the current available observations in finding the optimal parameter vector $\boldsymbol{\theta}$, may easily miss a more performing set of parameters. Indeed, a term promoting the *exploration* of the parameter space should be considered.

In the GLISp algorithm, an acquisition function is employed to balance exploitation *vs.* exploration when generating the new sample $\boldsymbol{\theta}_{N+1}$. The exploration function is defined by using the inverse distance weighting (IDW) function $z : \mathbb{R}^{n_\theta} \to \mathbb{R}$ defined as:

$$z(\boldsymbol{\theta}) = \begin{cases} 0 & \text{if } \boldsymbol{\theta} \in \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N\} \\ \tan^{-1}\left(\frac{1}{\sum_{i=1}^{N} w_i(\boldsymbol{\theta})}\right) & \text{otherwise,} \end{cases} \tag{52}$$

where $w_i(\boldsymbol{\theta}) = \frac{1}{d^2(\boldsymbol{\theta}, \boldsymbol{\theta}_i)}$. Clearly, $z(\boldsymbol{\theta}) = 0$ for all parameters already tested, and $z(\boldsymbol{\theta}) > 0$ in $\mathbb{R}^{n_\theta} \setminus \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N\}$. The arc tangent function in (52) avoids that $z(\boldsymbol{\theta})$ gets excessively large far away from all sampled points.

Then, given an exploration parameter $\delta \ge 0$, the *acquisition function* $a : \mathbb{R}^{n_\theta} \to \mathbb{R}$ is constructed as:

$$a(\boldsymbol{\theta}) = \frac{\hat{J}(\boldsymbol{\theta})}{\Delta\hat{J}} - \delta z(\boldsymbol{\theta}), \tag{53}$$

where

$$\Delta\hat{J} = \max_i \{\hat{J}(\boldsymbol{\theta}_i)\} - \min_i \{\hat{J}(\boldsymbol{\theta}_i)\}$$

is the range of the surrogate function on the samples in $\{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N\}$ and used in (53) as a normalization factor to simplify the choice of the exploration parameter δ.

Given a set $\{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N\}$ of samples and a vector $\mathbf{B}$ of preferences defined in (47), the next parameter set $\boldsymbol{\theta}_{N+1}$ to be tested is computed as the solution of the (non-convex) optimization problem:

$$\boldsymbol{\theta}_{N+1} = \arg\min_{\boldsymbol{\theta} \in \Theta} a(\boldsymbol{\theta}). \tag{54}$$

3.3. Pseudocode

The pseudocode of the employed preference-based optimization is detailed in Algorithm 1.

The PBO is interfaced with the proposed human–robot collaborative controller as depicted in Fig. 1.

Algorithm 1 PBO pseudocode

```
X ← [x_1, x_2, ..., x_N]
b_in ← [[x_1, x_2], [x_2, x_4], ..[x_h, x_k]]
b_eq ← [[x_2, x_3], ..., [x_z, x_y]]
my_pbo ← PBO(X, b_in, b_eq)
my_pbo.update()
for i →iterations do
    x_next ←my_pbo.run_optimization()
    if x_next better than x_best then
        eval ← −1
    else
        eval ← 1
    end if
    my_pbo.add_evaluation(x_next, eval)
    my_pbo.update()
end for
```

4. Experimental validation

In order to show the effectiveness of the proposed control scheme optimized by means of the PBO, the controller has been implemented on a Franka EMIKA panda robot, making use of its torque control (control frequency: 1000 Hz), for the execution of the target physical human–robot collaborative task. The experiment's purpose is twofold: to evaluate the performance of the proposed collaborative controller, comparing the achieved results w.r.t. the ones achieved by applying the controller in Roveda, Maskani, et al. (2020), and to evaluate the optimization process in terms of naturalness and easiness of use.

Remark 5. Considering the available high control frequency, all the controllers are designed for continuous time systems. For the discretization of the controllers, please refer to Roveda, Iannacci, and Tosatti (2018).

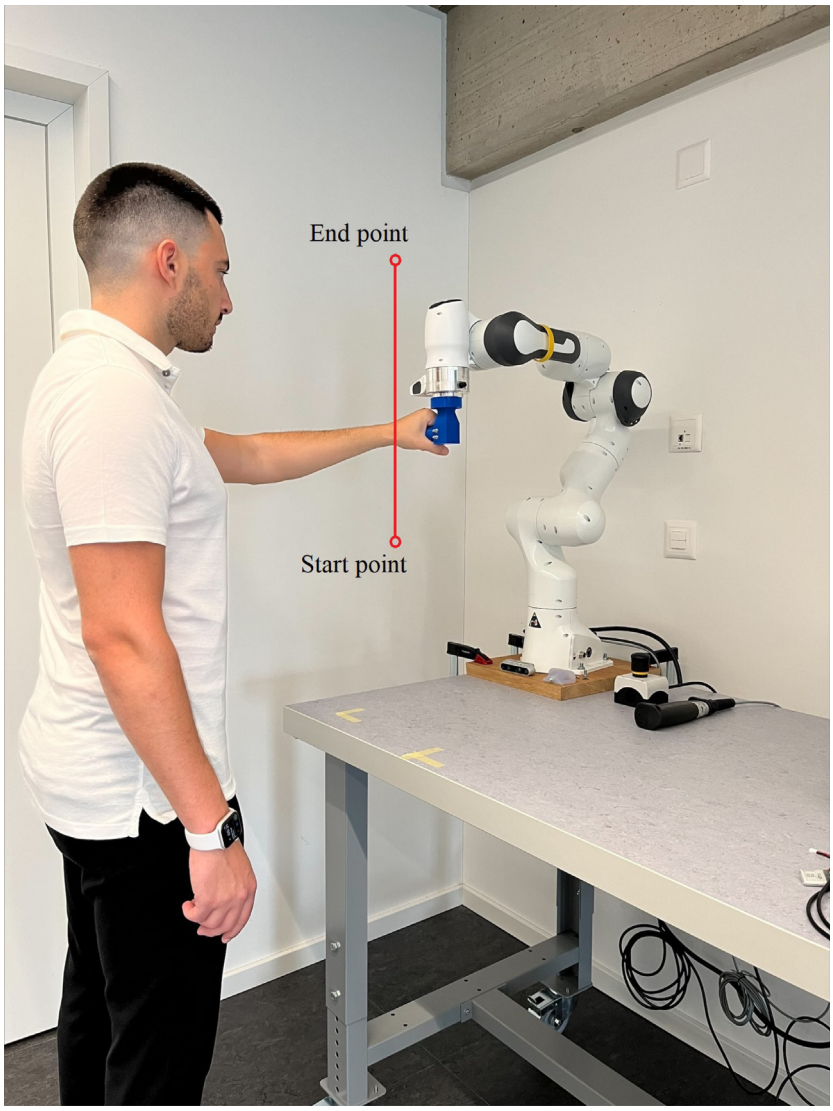


Fig. 2. Experimental tests performed having the volunteers in interaction with the employed Franka EMIKA panda robot.

4.1. Target pHRC task

The volunteers are supposed to operate the robot along the vertical direction z, where the proposed control approach and the one in Roveda, Maskani, et al. (2020) are activated. Two virtual walls along the z axis (that have been set at 0.7 m and 0.3 m) have been implemented in order to constrain the robot's motion. Each volunteer performs 2 up and down complete slow motions and 2 up and down complete fast motions between the defined virtual walls. A volunteer in interaction with the robot is shown in Fig. 2.

The proposed approach has been tuned by means of the PBO (*i.e.*, optimizing the Cartesian impedance control stiffness $K_{t,i}$ and damping $D_{t,i}$, and the LQR gain $q_{3,i}$) by each of the volunteers independently, so that each of them was able to customize its own controller. The collaborative performance of the optimized controller has been then compared with the ones obtained by the controller in Roveda, Maskani, et al. (2020).

Remark 6. The same physical human–robot collaboration task has been considered for the evaluation of the controllers as in Roveda et al. (2019) and Roveda, Maskani, et al. (2020).

4.2. Evaluation protocol

The proposed controller and the one in Roveda, Maskani, et al. (2020) have been evaluated on the basis of the qualitative evaluation framework used in Roveda et al. (2019), Roveda, Maskani, et al. (2020). Indeed, the physical collaboration performance for the two tested control methods have been assessed using a questionnaire at the end of the evaluation experiments for all the volunteers. The following metrics have been defined to address both physical human–robot interactions (pHRI) and cognitive human–robot interactions (cHRI) to completely address human expectations and task goals (Mizanoor Rahman & Ikeura, 2016):

- naturalness: human's overall likability, normalcy, ease of use, convenience, non-complexity in operation and collaboration;
- smoothness: whether the movement is smooth and characterized by a limited jerk;
- effort: the amount of effort, hardship, or endeavor required to achieve a performance level satisfying the mental, physical and temporal demand;

- motion: nature of object velocity and acceleration felt by the human (*i.e.*, whether the velocity, acceleration is low or high compared to human expectation);
- stability: presence/absence of oscillations, sudden inactivity of the system, and their effects on manipulation, object, system structure, and environment;
- detection of intention: whether the robot follows human intention in accelerating or deaccelerating the motion;
- performance: the overall performance, *e.g.*, lifting the object to the desired position within the specified time and attempting to avoid user-unfriendly events.

Each of the proposed metrics has a ranking between 0 (minimum ranking, very negative review) to 4 (maximum ranking, very positive review). The ratings have been assigned upon the completion of the experiment. Ratings are based on the subjective opinion of the participants, gained by the feeling with the controllers, and indicative mainly of their relative performance.

The optimization process (*i.e.*, making use of the PBO approach) has been evaluated in order to verify the naturalness and easiness of use of the proposed tuning procedure. The following criteria are considered for the proposed evaluation:

- user-friendliness: whether the training is intuitive or not (*i.e.*, if the necessary actions to train the robot are complex and/or too far from the ordinary way to use the robot);
- duration: the time required to achieve acceptable performance of the controller;
- satisfaction: whether the operator is able to achieve the target performance during the test.

Each of the proposed metrics has a ranking between 0 (minimum ranking) to 4 (maximum ranking).

4.3. Participants

The following volunteers have been involved in the experimental evaluation: 10 healthy subjects (9 males, 1 female, with mean age $= 31 \pm 6$ years) without any physical problems. Prior to the testing, all volunteers have been informed about the evaluation scenario and the testing procedure. Users are only trained in order to provide consistent feedback to the optimization algorithm. However, no prior objective is set for optimization purposes. In fact, each user might have his/her own intrinsic (and maybe unknown even to herself/himself) cost function (minimize the effort, fast and reactive collaboration, slow, robust, and stable collaboration, etc.) that is optimized based on the perceived collaboration.

4.4. Controller implementation

The code required to implement and run the proposed controller is available at the following GitHub repository: https://github.com/marcomaccarini/robotics/tree/main/lorenzo_ws.

The proposed controller has been implemented on a Franka EMIKA panda robot, exploiting the ROS framework. The implementation has been made on top of the available Franka torque control (based on libfranka functionalities), exploiting the Franka Control Interface (FCI), making it possible to send motor signals to the manipulator in real-time at 1 kHz. The interaction wrench is provided by the robot's internal sensors with a frequency of 1 kHz.

4.4.1. PVFC

The PVFC parameters have been imposed as it follows: $\bar{E} = 10$ J, $M_f = 1$ kg, $\gamma = 1$, and $\mathbf{h} = [25, 25, 15, 15, 100, 100, 400]^T$.

4.4.2. Cartesian impedance control

The rotational stiffness, damping, and inertia matrices have been imposed as follows: $\mathbf{K}_r = diag(100, 100, 100)$ Nm/rad, $\boldsymbol{\zeta}_r = diag(0.75, 0.75, 0.75)$, $\mathbf{M}_r = diag(10, 10, 10)$ Kgm^2. The translational stiffness $K_{t,i}$, damping $\zeta_{t,i}$, and mass $M_{t,i}$ parameters have been imposed equal to 1500 N/m, 1, and 10 kg, respectively, for $i = x$ and $i = y$. $M_{t,z} = 10$ kg has been imposed.

4.4.3. LQR

For the implementation of the LQR, the following parameters have been imposed as follows: human's arm stiffness $\mathbf{K}_h = diag(1002, 250, 320)$ N/m (as in Artemiadis et al., 2010), $\beta_{u,z} = 0.0001$ (as in Roveda & Piga, 2020), $U_{K_{h,z}} = \pm 0.5 K_{h,z}$ (as in Roveda & Piga, 2020), $q_{1,z} = 0.001$, and $q_{2,z} = 0.001$.

4.4.4. PBO

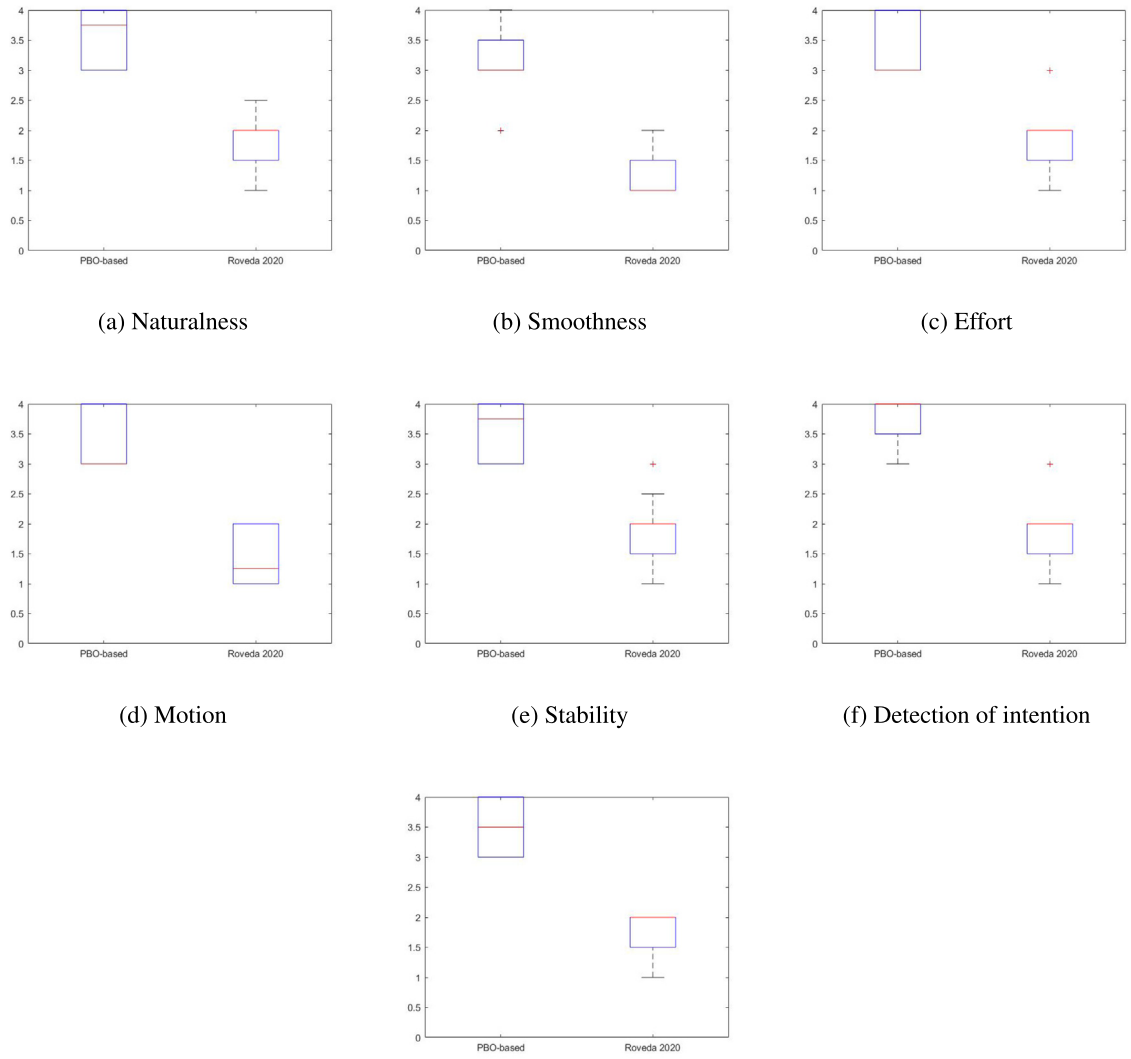
The following ranges of values have been considered for the optimization variables: $K_{t,z} \in [250, 1500]$ N/m, $\zeta_{t,z} \in [0.4, 2]$, and $q_{3,z} \in [0.02, 10]$. The definition of the parameters optimization range is important to avoid any issue for the controller (*e.g.*, stability) and can be defined based on the user's experience or similar works in the state of the art. 5 initial random sets of optimization parameters have been applied to each optimization (*i.e.*, for each volunteer) to initialize it. The parameter δ, regulating the exploration–exploitation behavior of the optimization algorithm (53), has been set to 3.

Remark 7. It has to be remarked that more parameters can be included in the optimization procedure if needed. In this paper, in fact, only human assistance-related gains are considered for optimization purposes. The other gains (such as the PVFC gains) are tuned based on state-of-the-art tuning or based on classical tuning procedures, such as Roveda, Forgione, and Piga (2020).

4.5. Experimental results

Based on the above-described considered task and evaluation protocols, the proposed optimized controller and the one in Roveda, Maskani, et al. (2020) have been tested. The results obtained considering the employed qualitative evaluation through questionnaires are shown in Fig. 3. As it can be highlighted, the proposed controller optimized by means of the PBO is capable to achieve better performance in comparison with the one in Roveda, Maskani, et al. (2020). In addition, it has to be underlined the capability of the proposed controller to deal with different users, *i.e.*, with different human arm stiffness values. In fact, each human has a different arm stiffness, and this value is adapted during the collaboration with the robot (Artemiadis et al., 2010). Indeed, the proposed robust LQR controller is capable to deal with this variation of the human arm stiffness, ensuring the stability of the interaction. In addition, the proposed optimization is capable to deal with different users' cost functions. In fact, each user has different (unknown and not predefined) objectives to establish a high-performance customized human–robot interaction. The proposed methodology is indeed able to catch the intrinsic objectives of the user, optimizing the perceived collaboration with the robot by means of qualitative feedback.

The controller in Roveda, Maskani, et al. (2020) explicitly aims to minimize the interaction force between the human and the robot (*i.e.*, minimizing the user effort). The optimization framework proposed in this paper, instead, aims to maximize the performance of the user's perceived collaboration, based on his/her specific (intrinsic and even unknown) cost function. In fact, the cost function employed by the operator (that is used to provide qualitative feedback to the algorithm) is not predefined, and the optimization is performed based on the user's sensations only. Thus, one expert operator might prefer a reactive robot behavior, resulting in an effortless collaboration, while a novel operator might prefer a slower and more stable collaboration, resulting



(a) Naturalness (b) Smoothness (c) Effort

(d) Motion (e) Stability (f) Detection of intention

(g) Overall performance

Fig. 3. Obtained results for the proposed optimized controller and for the one in Roveda, Maskani, et al. (2020).

in a higher applied force to operate the robot. Therefore, the achieved performance cannot be directly compared by means of quantitative force-based indexes but can be compared by means of perceived (*i.e.*, qualitative, as described above) effort and comfort.

The optimization procedure (*i.e.*, the PBO) has been evaluated to verify its applicability in real contexts (such as in the manufacturing industry). Fig. 4 shows the achieved results, from which it is possible to highlight the usefulness and the easiness of use of the PBO in the proposed context. The PBO is capable to catch the (even intrinsic and unknown) needs of the user to guide the optimization. Physical human–robot collaboration control benefits from this qualitative optimization approach, due to the subjective perception of the collaboration with the robot. In fact, the Satisfaction on the computed optimal control parameters achieved a high score, demonstrating the potential of the optimization in the context of the application.

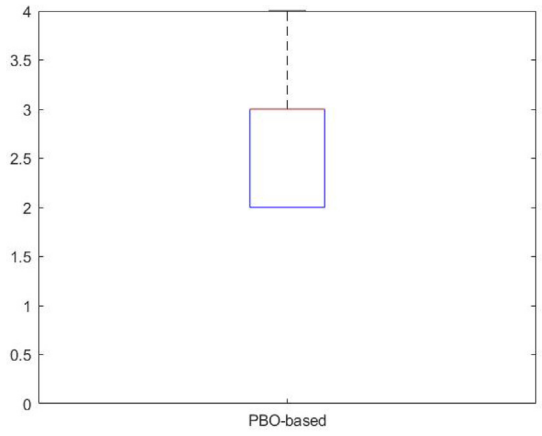
5. Validation use-case: collaborative assembly task

5.1. Task description

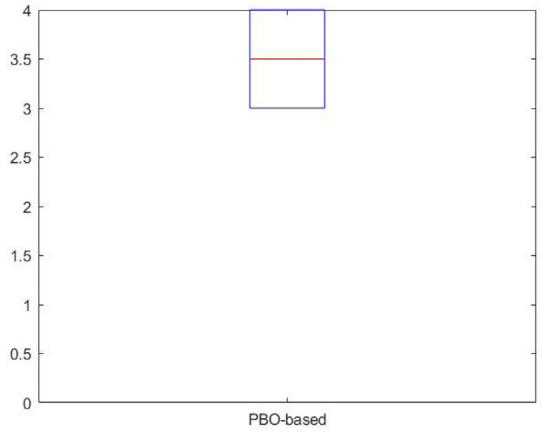
A more complex task involving all three translational DoFs has been taken into account to prove the capability of the proposed approach (controller + optimization procedure). A collaborative assembly task of a gear into its shaft is considered (Fig. 5). The same robot and controller implementation have been exploited as described in Section 4.

More in detail, the user has to operate the robot to perform the assembly. A video of the performed application during one optimization iteration is available here: https://youtu.be/OBf_CgGtCZQ. Following the proposed preference-based procedure, the controller parameters related to the provided assistance to the human are taken into consideration for optimization purposes (as described in Section 4.4). Given the nature of the proposed controller (*i.e.*, decoupled Cartesian impedance control with compensated gravity and friction effects), all three DoFs control parameters are optimized together in a single optimization (*i.e.*, one optimization assigning the same parameters to all the DoFs is performed). In fact, all the translational DoFs of the Cartesian impedance control are behaving the same (Siciliano et al., 2010).

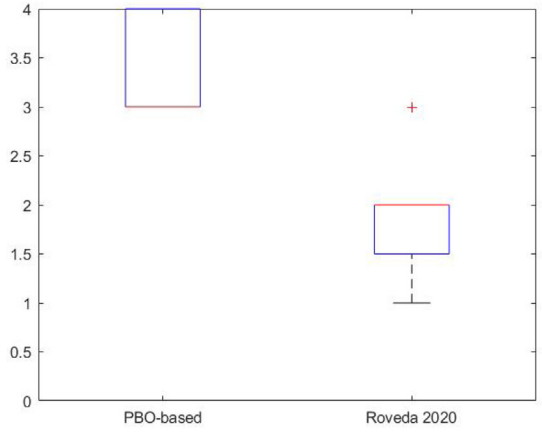
Remark 8. It has to be underlined that the three translational DoFs can be optimized even separately. Parallel optimizations might be run for each DoF independently. In this work, however, one optimization procedure has been considered due to the decoupled nature of the impedance control, allowing the DoFs to behave the same. This allows us to reduce the optimization time and the computational effort.



(a) Optimization time



(b) User friendliness



(c) Satisfaction on the computed optimal control parameters

Fig. 4. PBO process evaluation results.

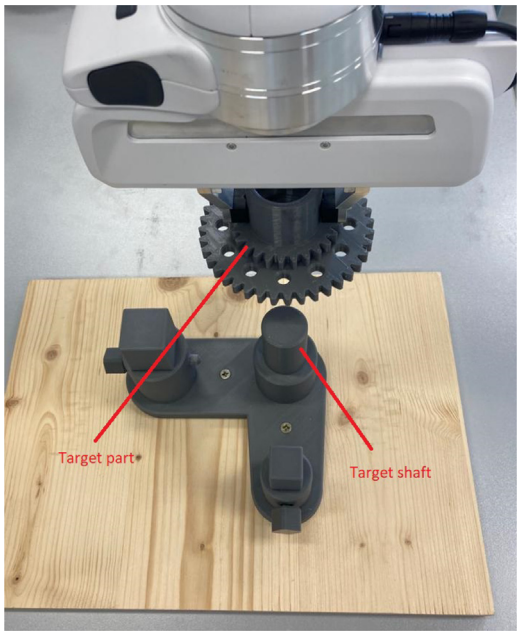


Fig. 5. Considered assembly task of gear to be inserted into its shaft.

5.2. Participants

The following volunteers have been involved in the experimental evaluation: 5 healthy subjects (4 males, 1 female, with mean age = 31 ± 5 years) without any physical problems. Prior to the testing, all volunteers have been informed about the evaluation scenario and the testing procedure. As in Section 4.3, users are only trained in order to provide consistent feedback to the optimization algorithm.

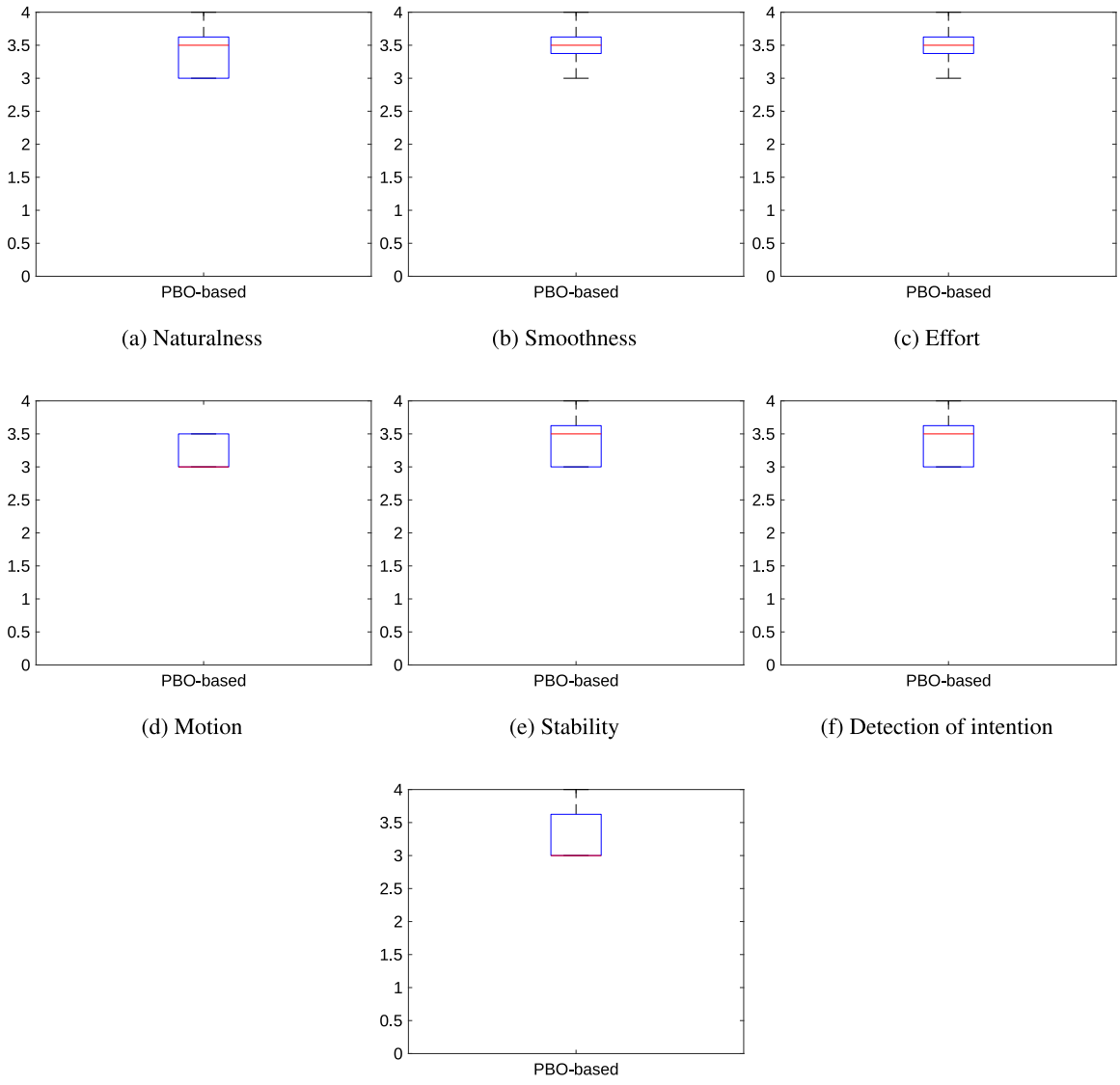
5.3. Experimental results

By means of the questionnaires described in Section 4.2, the achieved results for the optimization of the collaborative assembly task are reported in Figs. 6 and 7. Similarly to the results in Section 4.5, high scores are obtained for all the defined criteria, highlighting the performance of the proposed controller and user-centered optimization.

6. Conclusions and future work

This paper proposed a controller to enable a safe and robust physical human–robot collaboration while limiting the computational effort (to compute the control signals given the 1 kHz high control frequency) and providing a human-centric tuning of the control parameters providing assistance to maximize the collaboration for each specific user. The proposed controller has been evaluated by making use of a Franka EMIKA panda robot as a test platform considering a free-motion collaborative task (comparing the achieved performance with a controller previously developed by some of the authors) and an assembly task. Achieved results show the improved performance obtained by the proposed controller. In addition, the performance of the PBO for the tuning of the control parameters have been evaluated, highlighting the potentialities of the method.

Some limitations of the described control framework can be highlighted. W.r.t. the designed controller, it has to be noted that constant impedance control parameters are considered during the human–robot interaction task execution, which might limit the powerfulness of the control schema. W.r.t. the PBO for assistive gains optimization, such approach is only employing the qualitative feedback of the user (*i.e.*,



(a) Naturalness (b) Smoothness (c) Effort

(d) Motion (e) Stability (f) Detection of intention

(g) Overall performance

Fig. 6. Obtained results for the proposed optimized controller.

user's preferences). Such qualitative feedback has to be consistent among all the optimization iterations in order to provide a consistent optimized set of control parameters. One experienced issue during the optimization is that the user does not remember well his/her perception of the best set of parameters so far if too many iterations are passed, requiring an additional trail with such a set of parameters. Indeed, consistent judgments from the user are of critical importance.

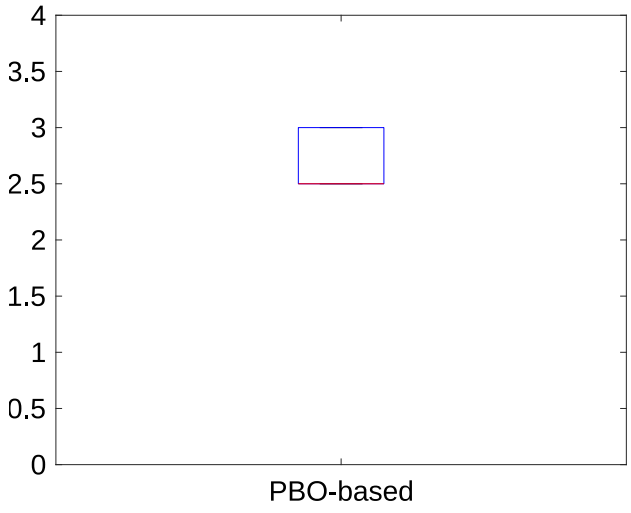
Future work will be devoted to integrating the features of the proposed control in the MBRL framework of Roveda, Maskani, et al. (2020) to provide user-based adaptability in real-time (*i.e.*, based on the operative conditions, adapting the control gains). In addition, ML-based controllers with stability guarantees are currently under development. Furthermore, ML techniques will be employed for the online estimation of the human's arm setpoint $\mathbf{x}_h^0$, stiffness $\mathbf{K}_h$, and intrinsic optimization cost function to improve the intention of motion detection to be used in the controller and the online human's arm dynamics modeling (*e.g.*, for the implementation of SDRE controllers Roveda & Piga, 2020). Finally, improvements related to the PBO algorithm are currently under consideration.

Declaration of competing interest

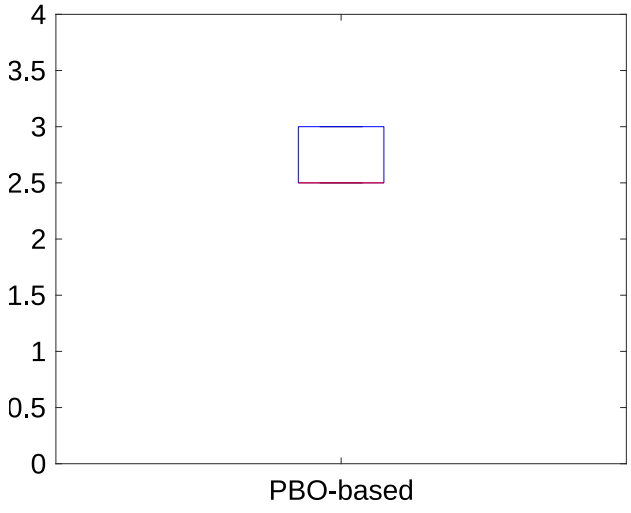
The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Loris Roveda reports financial support was provided by Hasler and EUROSTARS EUREKA.

Acknowledgments

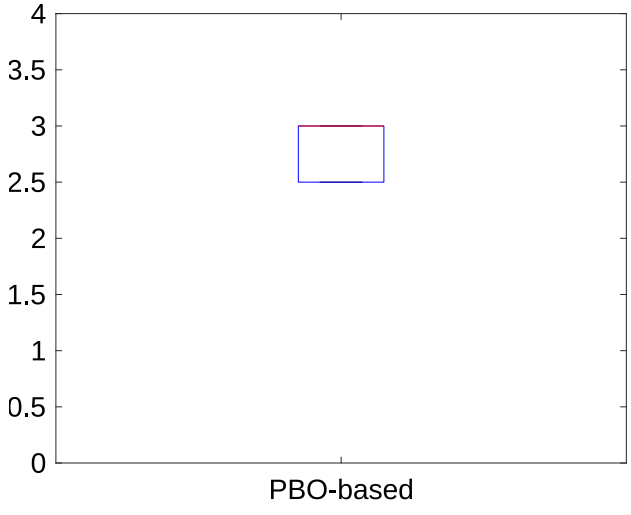
The work has been developed within the project *Preference-Based Optimization Tool Applied to Physical Human-Robot Collaboration* (RoboPref), funded by Hasler Foundation, Switzerland. The work has been developed within the project *ExoRescue*, funded from Eureka Eurostars under grant agreement n. E!115182.



(a) Optimization time



(b) User friendliness



(c) Satisfaction on the computed optimal control parameters

Fig. 7. PBO process evaluation results.

References

Artemiadis, P. K., Katsiaris, P. T., Liarokapis, M. V., & Kyriakopoulos, K. J. (2010). Human arm impedance: Characterization and modeling in 3d space. In *2010 IEEE/RSJ international conference on intelligent robots and systems* (pp. 3103–3108). IEEE.

Bemporad, A. (2020). Global optimization via inverse distance weighting and radial basis functions. (pp. 1–18). arxiv 8 (1).

Bemporad, A., Morari, M., Dua, V., & Pistikopoulos, E. N. (2002). The explicit linear quadratic regulator for constrained systems. *Automatica*, *38*(1), 3–20.

Bemporad, A., & Piga, D. (2021). Global optimization based on active preference learning with radial basis functions. *Machine Learning*, *110*(2), 417–448.

Busa-Fekete, R., Hüllermeier, E., & Mesaoudi-Paul, A. E. (2018). Preference-based online learning with dueling bandits: A survey. CoRR abs/1807.11398. URL http://arxiv.org/abs/1807.11398.

Caccavale, F., Natale, C., Siciliano, B., & Villani, L. (1999). Six-dof impedance control based on angle/axis representations. *IEEE Transactions on Robotics and Automation*, *15*(2), 289–300.

Choi, J., Dance, C., Kim, J.-e., Park, K.-s., Han, J., Seo, J., et al. (2020). Fast adaptation of deep reinforcement learning-based navigation skills to human preference. In *2020 IEEE international conference on robotics and automation* (pp. 3363–3370). IEEE.

Çimen, T. (2008). State-dependent Riccati equation (SDRE) control: A survey. *IFAC Proceedings Volumes*, *41*(2), 3761–3775.

Colgate, E., & Hogan, N. (1989). An analysis of contact instability in terms of passive physical equivalents. In *Proceedings, 1989 international conference on robotics and automation* (pp. 404–409). IEEE.

Cremer, S., Das, S. K., Wijayasinghe, I. B., Popa, D. O., & Lewis, F. L. (2019). Model-free online neuroadaptive controller with intent estimation for physical human–robot interaction. *IEEE Transactions on Robotics*, *36*(1), 240–253.

Demir, K. A., Döven, G., & Sezen, B. (2019). Industry 5.0 and human-robot co-working. *Procedia Computer Science*, *158*, 688–695.

Dimeas, F., & Aspragathos, N. (2015). Reinforcement learning of variable admittance control for human-robot co-manipulation. In *2015 IEEE/RSJ international conference on intelligent robots and systems* (pp. 1011–1016). IEEE.

Fisher, J., & Bhattacharya, R. (2009). Linear quadratic regulation of systems with stochastic parameter uncertainties. *Automatica*, *45*(12), 2831–2841.

Fligge, N., McIntyre, J., & van der Smagt, P. (2012). Minimum jerk for human catching movements in 3D. In *2012 4th IEEE RAS & EMBS international conference on biomedical robotics and biomechatronics* (pp. 581–586). IEEE.

Franceschi, P., Pedrocchi, N., & Beschi, M. (2022). Adaptive impedance controller for human-robot arbitration based on cooperative differential game theory. In *2022 international conference on robotics and automation* (pp. 7881–7887). IEEE.

Freeman, R. A., & Kokotovic, P. V. (1996). Inverse optimality in robust stabilization. *SIAM Journal on Control and Optimization*, *34*(4), 1365–1391.

Fu, L., & Zhao, J. (2020). Maxwell-model-based compliance control for human–robot friendly interaction. *IEEE Transactions on Cognitive and Developmental Systems*, *13*(1), 118–131.

Galin, R. R., & Meshcheryakov, R. V. (2020). Human-robot interaction efficiency and human-robot collaboration. In *Robotics: Industry 4.0 issues & new intelligent control paradigms* (pp. 55–63). Springer.

Gaz, C., Magrini, E., & De Luca, A. (2018). A model-based residual approach for human-robot collaboration during manual polishing operations. *Mechatronics*, *55*, 234–247.

Grüne, L., & Pannek, J. (2017). Nonlinear model predictive control. In *Nonlinear model predictive control* (pp. 45–69). Springer.

He, W., Xue, C., Yu, X., Li, Z., & Yang, C. (2020). Admittance-based controller design for physical human–robot interaction in the constrained task space. *IEEE Transactions on Automation Science and Engineering*, *17*(4), 1937–1949.

Hentout, A., Aouache, M., Maoudj, A., & Akli, I. (2019). Human–robot interaction in industrial collaborative robotics: A literature review of the decade 2008–2017. *Advanced Robotics*, *33*(15–16), 764–799.

Hogan, N. (1984). Impedance control: An approach to manipulation. In *1984 American control conference* (pp. 304–313). IEEE.

Kennedy, J. (2010). Particle swarm optimization. In *Encyclopedia of machine learning* (pp. 760–766). Boston, MA: Springer US, http://dx.doi.org/10.1007/978-0-387-30164-8_630.

Khan, S. G., Herrmann, G., Al Grafi, M., Pipe, T., & Melhuish, C. (2014). Compliance control and human–robot interaction: Part 1—survey. *International Journal of Humanoid Robotics*, *11*(03), Article 1430001.

Khoramshahi, M., & Billard, A. (2020). A dynamical system approach for detection and reaction to human guidance in physical human–robot interaction. *Autonomous Robots*, *44*(8), 1411–1429.

Kim, W., Peternel, L., Lorenzini, M., Babič, J., & Ajoudani, A. (2021). A human-robot collaboration framework for improving ergonomics during dexterous operation of power tools. *Robotics and Computer-Integrated Manufacturing*, *68*, Article 102084.

Kronander, K., & Billard, A. (2012). Online learning of varying stiffness through physical human-robot interaction. In *2012 IEEE international conference on robotics and automation* (pp. 1842–1849). Ieee.

Kukker, A., & Sharma, R. (2021). Stochastic genetic algorithm-assisted fuzzy Q-learning for robotic manipulators. *Arabian Journal for Science and Engineering*, 1–13.

Lee, K., Smith, L., Dragan, A., & Abbeel, P. (2021). B-pref: Benchmarking preference-based reinforcement learning. http://dx.doi.org/10.48550/ARXIV.2111.03026, URL https://arxiv.org/abs/2111.03026.

Li, P. Y., & Horowitz, R. (1999). Passive velocity field control of mechanical manipulators. *IEEE Transactions on Robotics and Automation*, *15*(4), 751–763.

Li, P. Y., & Horowitz, R. (2001a). Passive velocity field control (PVFC). Part I. Geometry and robustness. *IEEE Transactions on Automatic Control*, *46*(9), 1346–1359.

Li, P. Y., & Horowitz, R. (2001b). Passive velocity field control (PVFC). Part II. Application to contour following. *IEEE Transactions on Automatic Control*, *46*(9), 1360–1371.

Li, C., Zhang, Z., Xia, G., Xie, X., & Zhu, Q. (2018). Efficient force control learning system for industrial robots based on variable impedance control. *Sensors*, *18*(8), 2539.

Liang, P., Yang, C., Wang, N., Li, Z., Li, R., & Burdet, E. (2014). Implementation and test of human-operated and human-like adaptive impedance controls on baxter robot. In *Conference towards autonomous robotic systems* (pp. 109–119). Springer.

Losey, D. P., McDonald, C. G., Battaglia, E., & O'Malley, M. K. (2018). A review of intent detection, arbitration, and communication aspects of shared control for physical human–robot interaction. *Applied Mechanics Reviews*, *70*(1).

Ma, W. -L., Kolathaya, S., Ambrose, E. R., Hubicki, C. M., & Ames, A. D. (2017). Bipedal robotic running with DURUS-2D: Bridging the gap between theory and experiment. In *Proceedings of the 20th international conference on hybrid systems: Computation and control* (pp. 265–274).

Magrini, E., & De Luca, A. (2016). Hybrid force/velocity control for physical human-robot collaboration tasks. In *2016 IEEE/RSJ international conference on intelligent robots and systems* (pp. 857–863). IEEE.

Makris, S. (2021). *Cooperating robots for flexible manufacturing*. Springer.

Mariotti, E., Magrini, E., & De Luca, A. (2019). Admittance control for human-robot interaction using an industrial robot equipped with a F/T sensor. In *2019 international conference on robotics and automation* (pp. 6130–6136). IEEE.

Martinez, A., Lawson, B., Durrough, C., & Goldfarb, M. (2018). A velocity-field-based controller for assisting leg movement during walking with a bilateral hip and knee lower limb exoskeleton. *IEEE Transactions on Robotics*, *35*(2), 307–316.

McDonald, D. B., Grantham, W. J., Tabor, W. L., & Murphy, M. J. (2007). Global and local optimization using radial basis function response surface models. *Applied Mathematical Modelling*, *31*(10), 2095–2110. http://dx.doi.org/10.1016/j.apm.2006.08.008, URL https://www.sciencedirect.com/science/article/pii/S0307904X06002009.

Medina, J. R., Börner, H., Endo, S., & Hirche, S. (2019). Impedance-based gaussian processes for modeling human motor behavior in physical and non-physical interaction. *IEEE Transactions on Biomedical Engineering*, *66*(9), 2499–2511.

Mizanoor Rahman, S., & Ikeura, R. (2016). Cognition-based control and optimization algorithms for optimizing human-robot interactions in power-assisted object manipulation.. *Journal of Information Science & Engineering*, *32*(5).

Mobasser, F., & Hashtrudi-Zaad, K. (2006). A method for online estimation of human arm dynamics. In *2006 international conference of the IEEE engineering in medicine and biology society* (pp. 2412–2416). IEEE.

Palan, M., Landolfi, N. C., Shevchuk, G., & Sadigh, D. (2019). Learning reward functions by integrating human demonstrations and preferences. arXiv preprint arXiv:1906.08928.

Peternel, L., Tsagarakis, N., & Ajoudani, A. (2017). A human–robot co-manipulation approach based on human sensorimotor information. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(7), 811–822.

Rodríguez-Guerra, D., Sorrosal, G., Cabanes, I., & Calleja, C. (2021). Human-robot interaction review: Challenges and solutions for modern industrial environments. *IEEE Access*, *9*, 108557–108578.

Roveda, L. (2018). A user-intention based adaptive manual guidance with force-tracking capabilities applied to walk-through programming for industrial robots. In *2018 15th international conference on ubiquitous robots* (pp. 369–376). IEEE.

Roveda, L., Bussolan, A., Braghin, F., & Piga, D. (2022). Robot joint friction compensation learning enhanced by 6D virtual sensor. *International Journal of Robust and Nonlinear Control*.

Roveda, L., Castaman, N., Franceschi, P., Ghidoni, S., & Pedrocchi, N. (2020). A control framework definition to overcome position/interaction dynamics uncertainties in force-controlled tasks. In *2020 IEEE international conference on robotics and automation* (pp. 6819–6825). IEEE.

Roveda, L., Forgione, M., & Piga, D. (2020). Robot control parameters auto-tuning in trajectory tracking applications. *Control Engineering Practice*, *101*, Article 104488.

Roveda, L., Haghshenas, S., Caimmi, M., Pedrocchi, N., & Molinari Tosatti, L. (2019). Assisting operators in heavy industrial tasks: On the design of an optimized cooperative impedance fuzzy-controller with embedded safety rules. *Frontiers in Robotics and AI*, 75.

Roveda, L., Haghshenas, S., Prini, A., Dinon, T., Pedrocchi, N., Braghin, F., et al. (2018). Fuzzy impedance control for enhancing capabilities of humans in onerous tasks execution. In *2018 15th international conference on ubiquitous robots* (pp. 406–411). IEEE.

Roveda, L., Iannacci, N., & Tosatti, L. M. (2018). Discrete-time formulation for optimal impact control in interaction tasks. *Journal of Intelligent and Robotic Systems*, *90*(3–4), 407–417.

Roveda, L., Iannacci, N., Vicentini, F., Pedrocchi, N., Braghin, F., & Tosatti, L. M. (2015). Optimal impedance force-tracking control design with impact formulation for interaction tasks. *IEEE Robotics and Automation Letters*, *1*(1), 130–136.

Roveda, L., Maggioni, B., Marescotti, E., Shahid, A. A., Zanchettin, A. M., Bemporad, A., et al. (2021). Pairwise preferences-based optimization of a path-based velocity planner in robotic sealing tasks. *IEEE Robotics and Automation Letters*, *6*(4), 6632–6639.

Roveda, L., Maskani, J., Franceschi, P., Abdi, A., Braghin, F., Molinari Tosatti, L., et al. (2020). Model-based reinforcement learning variable impedance control for human-robot collaboration. *Journal of Intelligent and Robotic Systems*, *100*(2), 417–433.

Roveda, L., & Piga, D. (2020). Robust state dependent riccati equation variable impedance control for robotic force-tracking tasks. *International Journal of Intelligent Robotics and Applications*, *4*(4), 507–519.

Sciavicco, L., & Siciliano, B. (2012). *Modelling and control of robot manipulators*. Springer Science & Business Media.

Siciliano, B., Sciavicco, L., Villani, L., & Oriolo, G. (2010). *Robotics: Modelling, planning and control*. Springer Science & Business Media.

Tee, K. P., Burdet, E., Chew, C. -M., & Milner, T. E. (2004). A model of force and impedance in human arm movements. *Biological Cybernetics*, *90*(5), 368–375.

Tian, R., Sun, L., Bajcsy, A., Tomizuka, M., & Dragan, A. D. (2022). Safety assurances for human-robot interaction via confidence-aware game-theoretic human models. In *2022 international conference on robotics and automation* (pp. 11229–11235). IEEE.

Vaz, A. I. F., & Vicente, L. N. (2007). A particle swarm pattern search method for bound constrained global optimization. *Journal of Global Optimization*, *39*(2), 197–219. http://dx.doi.org/10.1007/s10898-007-9133-5.

Vicentini, F., Pedrocchi, N., Beschi, M., Giussani, M., Iannacci, N., Magnoni, P., et al. (2020). PIROS: Cooperative, safe and reconfigurable robotic companion for CNC pallets load/unload stations. In *Bringing innovative robotic technologies from research labs to industrial end-users* (pp. 57–96). Springer.

Yang, C., Zeng, C., Fang, C., He, W., & Li, Z. (2018). A DMPs-based framework for robot learning and generalization of humanlike variable impedance skills. *IEEE/ASME Transactions on Mechatronics*, *23*(3), 1193–1203.

Yue, Y., Broder, J., Kleinberg, R., & Joachims, T. (2012). The K-armed dueling bandits problem. *Journal of Computer and System Sciences*, *78*(5), 1538–1556. http://dx.doi.org/10.1016/j.jcss.2011.12.028, URL https://www.sciencedirect.com/science/article/pii/S0022000012000281, JCSS Special Issue: Cloud Computing 2011.

Zacharaki, A., Kostavelis, I., Gasteratos, A., & Dokas, I. (2020). Safety bounds in human robot interaction: A survey. *Safety Science*, *127*, Article 104667.

Zhang, H., Li, S., & Zheng, Y. (2020). Q-learning-based model predictive control for nonlinear continuous-time systems. *Industrial and Engineering Chemistry Research*, *59*(40), 17987–17999.

Zhu, M., Bemporad, A., & Piga, D. (2020). Preference-based MPC calibration. arXiv: 2003.11294.